MISFIRE FAULT DETECTION IN SPARK IGNITION ENGINE USING HYBRID MODEL

by

Muddassar Abbas Rizvi
Reg. No. PE073002

Thesis submitted to the
Electronic Engineering Department
in partial fulfillment of the requirements for the degree of
Ph.D. IN ELECTRONIC ENGINEERING

Faculty of Engineering and Applied Sciences
Mohammad Ali Jinnah University
Islamabad
August, 2010

To

My Family and Teachers

whose help and efforts enabled me to pursue

my

Education

ACKNOWLEDGMENT

This work has been performed at Electronic Engineering Department, Mohammad Ali Jinnah University; Islamabad Campus. The work is financially supported by Higher Education Commission (HEC), Pakistan and ICT R&D Fund, Pakistan. First of all the author is thankful to Allah the almighty whose benevolence rendered him to achieve the prestigious goals of study. The author highly acknowledged the help of his family on account of their support during the study period. Author is also grateful to his advisor *Prof. Dr. Aamer Iqbal Bhatti* and all his colleagues for their help and contributions on account of research and development being carried out during the course of this work. Author is also thankful to Dr.Abdul Qayyum Khan for his help in the foreign evaluation of thesis.

Author is also very thankful to *Prof. Dr. Elias G. Strangas*, Associate professor, Department of Electrical and Computer Engineering, Michigan State University, East Lansing, USA and *Prof. Dr. Nadhir Messai*, Université de Reims Champagne-Ardenne UFR SEN, CReSTIC, Moulin de la Housse BP 1039, who spared their valuable time to review the work presented in this thesis and provided their valuable comments for the improvement of work.

Author also acknowledges the effort of the development team working with him to carry out the social aspect of project that has contributed to the professional uplift of some of the automotive mechanics working in the local market.

ABSTRACT

Automotive industry has added the self-diagnostic features in vehicles to improve the reliability of vehicle. Research is being carried out to predict the faults that are going to occur in near future by the analysis of current values of vehicle variables. The presented work stressed on the application of Markov chains for the early detection of misfire fault in spark ignition engines. To define the states of Markov chains a novel hybrid model is presented to represent SI engine under steady state conditions.

A survey of existing mathematical models of SI engine is provided. The hybrid model of SI engine was not widely studied area in the past. The proposed hybrid model with both continuous and discrete states is described in details. The basic assumption of modeling is that the cylinder contributing engine power is the basic active sub-component that provides power for useful work as well as to other cylinders that need power for compression, suction or exhaust. The cylinder providing power is considered as the active cylinder. The active cylinder is switched periodically in a cyclic manner.

The continuous states of hybrid model are defined by considering each cylinder of SI engine as the sub-systems of hybrid model. The switching of active cylinder is considered as discrete state of hybrid model. The model is simulated to study the crankshaft speed fluctuations observed in SI engine. The simulation results are then verified experimentally on 1300 cc engine of a production vehicle from Honda by acquiring data using Data Acquisition Cards of National Instrument Inc. The properties of presented model are then studied and some results are established for onward stochastic analysis.

The crankshaft speed fluctuation signal is analyzed using the properties of the proposed model and it is established that the peak values of observed speed during an ignition cycle is Gaussian and Markov. The peak value of crankshaft speed observed in each ignition cycle is associated with one of the cylinders or sub-systems. In this way four possible states are identified where i^{th} state correspond to the peak value of crankshaft speed associated with i^{th} sub-system of hybrid model. It is assumed that all states are equally probable when engine is healthy and that the fault would bias one of the states. The proposed novel fault detection algorithm identifies the biasing of a state by the calculation of Limiting State Probability of Markov Chains to indicate the fault.

The data for both healthy and faulty engine condition is generated using hybrid model and analyzed using proposed fault detection method. The algorithm is finally verified experimentally by acquiring data from SI engine both under no fault condition and faulty condition and analyzing it for the existence of fault.

The correctness of fault predicted by algorithm is mathematically analyzed using analysis similar to ROC analysis. In error analysis the fault is predicted using proposed algorithm and compared with the data observed experimentally to study the false positive events. The plot of analysis demonstrates the affectivity of algorithm.

TABLE OF CONTENTS

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

SI .... Spark Ignition
MIL .... Malfunction Indication Lamp
OBD .... On Board Diagnostics
CAN .... Control Area Network
FDI .... Fault Detection and Isolation
FDII .... Fault Detection, Isolation and Identification
MVM .... Mean Value Model
PDF .... Probability Density Function
TDC .... Top Dead Center
BDC .... Bottom Dead Center
PV Diagram .... Pressure Volume Diagram
MAP .... Manifold Absolute Pressure
MAF .... Manifold Absolute Flow
IFAC .... International Federation of Automation and Control
CBM .... Condition Based Monitoring
LMI .... Linear Matrix Inequality
FDF .... Fault Detection Filters
ARMA .... Auto Regressive and Moving Average
PCA .... Principal Component Analysis
ANN .... Artificial Neural Network
NN .... Neural Network
HMM .... Hidden Markov Model
DEM .... Discrete Event Model
ODE .... Ordinary Differential Equation

Chapter 1

INTRODUCTION

1.1 Overview

The development of efficient algorithms for "Fault Detection and Isolation (FDI) in Spark Ignition (SI) Engine" is a major research area in the field of automotive research. This research proposal is oriented towards "***Detection of Misfire Fault in SI Engine using a novel Hybrid Model to represent SI engine***". In this thesis the study contains following two novelties:

- ***A novel hybrid model of SI engine***
- ***A novel misfire fault detection technique on the basis of Markov Chains***

in this research proposal *Misfire Fault* is defined as per definition of California Environmental protection agency, Air Resources Board (CARB) (2009, pp: 117) i.e. "*lack of combustion in the cylinder due to absence of spark, poor fuel metering, poor compression, or any other cause. Lack of combustion events in non-active cylinders due to default fuel shut-off or cylinder deactivation strategies are however not considered as misfire event*."

This research field originated with the awareness of pollution hazard caused by the accumulation of exhaust emission of billions of vehicles running on roads around the world. Engine faults not only aggravate the emission problems but also result in fuel wastage and passenger discomfort due to non-smooth vehicle movement and even vehicle failure. Considering the environmental hazards on account of large number of vehicles on road, the permissible limits were legally defined for exhaust emissions in a number of countries. As an example, CARB originally adopted the light and medium duty vehicle OBD regulation (OBD-II) in 1989 for vehicles of 1996 and later models. The vehicle manufacturers invested in the development of ***Electronic Fuel Injection (EFI)*** and electronic controllers to ensure the optimal burning of fuel and minimize the exhaust emissions in SI engine through optimal design of engine controller. As per H. Holzmann etal, (1999, pp. 1014-1019), the basic objectives considered in EFI engine design include:

- Vehicles with better fuel efficiency
- Vehicles with less exhaust emissions
- Vehicles with better safety aspects
- Better availability of vehicle
- Passenger comfort

The engine controllers are designed to work under the designed conditions. The occurrence of certain engine faults however change the engine system in such a manner that even electronic controller cannot ensure the achievement of desired performance objectives. Research interest is therefore developed in the timely detection and identification of engine faults. The interest in the field of "Fault Detection, Isolation and Identification in SI engine" is further augmented due to the depleting energy resources and increasing fuel prices.

To meet the major challenges of twenty first century, the computational power of microprocessors and microcontrollers is utilized in automotives to achieve the performance objectives. The decreasing prices of electronics equipment makes it possible to install sensors in vehicles that measure the quantity of air sucked in engine cylinder and spray desired amount of fuel in it that would be burnt completely by the sucked air. Similarly sensors and actuators are installed in exhaust system to control the exhaust emission. The control signals are provided to the actuator through an ***Electronic Control Unit*** (***ECU***). The ECU is a microcontroller based electronic circuit that controls the quantity of fuel injected in the cylinder and the spark position. The EFI control resulted in improved vehicle performance.

Even in the presence of all these sensors and actuators, the performance of vehicle would become sub-optimal when a fault occurs in some vehicle system. To ensure the desired performance of vehicle, the life of different components of vehicle is defined and it is recommended that the component be replaced after the end of that life. To achieve the objective of maintaining optimal performance of vehicle, it is necessary to observe the ***preventive maintenance*** schedule of vehicle e.g. to change the spark plugs after the vehicle travel some specified distance. The vehicle manufacturers suggest fairly pessimistic estimates of life different components and if the components are not replaced they may continue to work properly. This resulted in increase in maintenance cost of vehicle.

In order to optimize the maintenance cost of vehicles, it is necessary to shift the *maintenance strategy of vehicle from Preventive Maintenance (PM) to* **Condition Base Maintenance** (**CBM).** In CBM, the performance of a component is analyzed by analyzing the system output. The component is changed only when the performance of system is significantly deteriorated due to its fault. The average lifetime of the components is therefore increased by CBM. For implementing strategies based on CBM it is necessary to develop **fault detection techniques** capable of detecting the incipient faults. The basic philosophy of fault detection methods proposed for automotive industry was adopted from the methods being used in **Safety Critical Systems** like aircraft. These methods were then tailored for application in automotive area. The algorithm can be executed in a microcontroller of vehicle ECU.

As a first step, all the major faults that may affect the vehicle performance are monitored by ECU. The objective of detection and annunciation of fault is achieved by continuous monitoring of different sensors to detect the faults present in systems and indicates the fault through a ***Malfunction Indication Lamp*** (***MIL***) in the vehicle that provides a visual indication of fault to the driver. Appropriate fault codes are generated corresponding to each fault. These fault codes can be observed by connecting a diagnostic tester with the ***On Board Diagnostic*** (***OBD***) jack provided in a vehicle. The diagnostic tester communicates with the vehicle through standard communication protocols like ISO-9141-2 or ***Control Area Network*** (***CAN***) protocol [Protocol Document ISO 9141-2, 2000, Protocol Document SAE J-1979, 2002].

Some basic shortcomings observed in these approaches include:

- A malfunction is indicated in vehicle only when the fault is increased to such an extent that failure of component could be detected. Before the fault become apparent, the system however continues to operate in sub-optimal mode.
- The suboptimal operation of vehicle before MIL indication comes, result in decrease in fuel economy and excessive impermissible exhaust emissions.

During the last decades of twentieth century, research was initiated on the fault detection and isolation in vehicular systems with special emphasis on early detection of faults. The development of efficient algorithms for "***Fault Detection and Isolation in Spark Ignition (SI) engine***" is still a major research area in the field of automotives.

The research interest in this field was further grown when the concept of autonomous vehicles was introduced to enhance the safety features in the new generation vehicles. These vehicles ensure lane discipline by automatic steering and safe distance from other vehicles and objects by automatic braking. To ensure automatic steering and automatic braking, actuators (DC motors) were placed in vehicles which receive signals through a controller implemented in a microcontroller in ECU. The philosophy was in general termed as ***Drive By Wire*** (***DBW***). The two major components of DBW philosophy are ***Brake By Wire*** (***BBW***) and ***Steer By Wire*** (***SBW***). Both these subsystems are safety critical systems for which ***Fault Tolerant Control (FTC)*** is needed. BBW was first introduced in Mercedes Benz SL series in 2001-02. Under manual control, the system received input from the pressing of paddle from driver and a controller implemented in a microcontroller unit actuates the braking system of vehicle. The controller may however itself actuate the brakes after sensing some hazardous condition. The system was however removed after a few years due to problems [Huaqun G., 2009, pp: 13]. To ensure a high reliability for these systems, automatic fault detection and fault tolerance capability is needed.

Only a few features of modern vehicle that are implemented through a microcontroller are described above. However the implementation of these features is difficult in a single microcontroller. To implement all the features presented in a modern vehicle, the higher end vehicles use as many as 70 different microcontrollers. The vehicles of modern generation therefore represent a network of computer systems in which a number of microcontrollers share their data with each other and implement a number of control loops, communication protocols and diagnostic services etc. In these vehicles fault detection can be carried out on a microcontroller.

To clearly define the problem and discusses the solution to the problem, the basic terminology of SI engine would be required. It is appropriate to provide a brief overview of mechanical construction of SI engine and discuss the basic thermodynamics principles working behind an SI engine.

1.2 Introduction to SI Engine

The most common application of a spark ignition (SI) engine is in an automotive. The automotive engines consist of three, four or six cylinders. Each cylinder is equipped with two ports one connected to the input manifold and other port is connected to the output manifold. A piston moves inside the cylinder which is connected to shaft through a crankshaft mechanism. Figure 1.1 indicates a single engine cylinder indicating Top Dead Center (TDC) and Bottom Dead Center (BDC). In a four cylinder SI engine, four cylinders are coupled on a common shaft in such a way that power produced by each of the cylinders sum up to produce the total power.

SI engines are four stroke engines that work on the basis of Otto cycle. An Otto Cycle

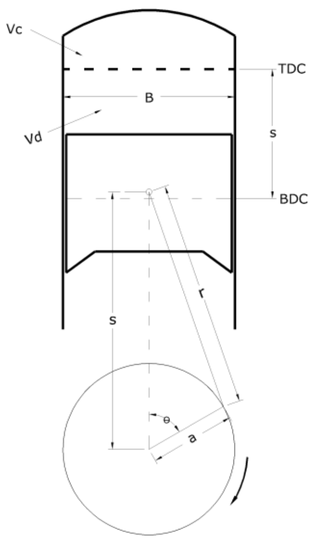


Figure 1.1: Line Diagram of Cylinder of SI Engine

is completed in four independent strokes of piston called:

- Suction stroke
- Compression stroke
- Power stroke
- Exhaust stroke.

Suction Stroke starts with the piston at Top Dead Center (TDC) and is characterized as a constant-pressure process. The inlet port of engine is opened and air is sucked from the intake manifold in the cylinder as the piston moves from TDC to Bottom Dead Center (BDC).

Compression stroke is an isentropic compression and is started when the piston is at BDC and ends when piston reaches TDC. The temperature within the cylinder is increased substantially due to compressive heating.

The addition of heat energy is assumed to be a constant-volume heat input process at TDC. In real engine this occurs at close to constant-volume conditions as heat input is started when piston is slightly before TDC and ends with piston slightly after TDC. During this process a large amount of energy is added to the air within the cylinder. This energy raises the temperature of the air to very high values.

The ignition of air fuel mixture results in the generation of very high pressure and enthalpy values within the system at TDC. This high pressure inside the cylinder results in a power stroke. During power stroke some of the energy is converted to work causing the piston to move from TDC to BDC. Due to the rise of temperature in cylinder some energy is transferred to cooling system through engine cooling ducts.

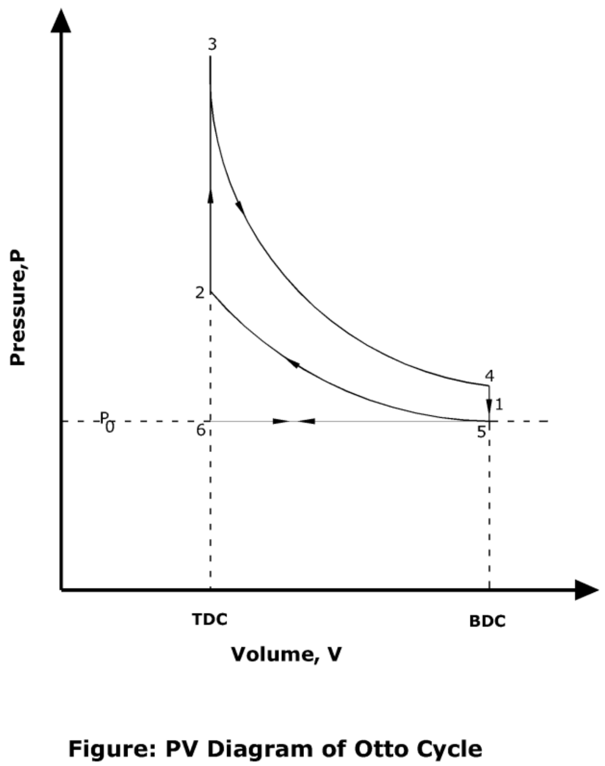


Figure 1.2 : PV Diagram of Otto Cycle

The rest of the energy remains in the cylinder in the form of hot gases.

Exhaust stroke occurs when piston is at BDC. The exhaust port opens and as the piston moves from BDC to TDC, the hot gases are pushed out of the cylinder to the atmosphere through engine exhaust. A complete ideal Otto Cycle is represented by a closed curve on a PV diagram shown in Figure 1.2.

During each stroke the crankshaft is rotated by 180˚. The ignition cycle of an SI engine is therefore completed during an angular movement of 720˚ or two complete rotations of crankshaft. The four cylinders of an engine are 180˚ out of phase from each other and the nature of stroke in all the cylinders is different at any particular instant e.g. if one cylinder is sucking air, some other cylinder would be compressing the air sucked by it in the previous cycle and one of the cylinder would be generating the power and the remaining cylinder would be exhausting the burnt gases. This arrangement of cylinder ensures that one of the cylinders would be generating power

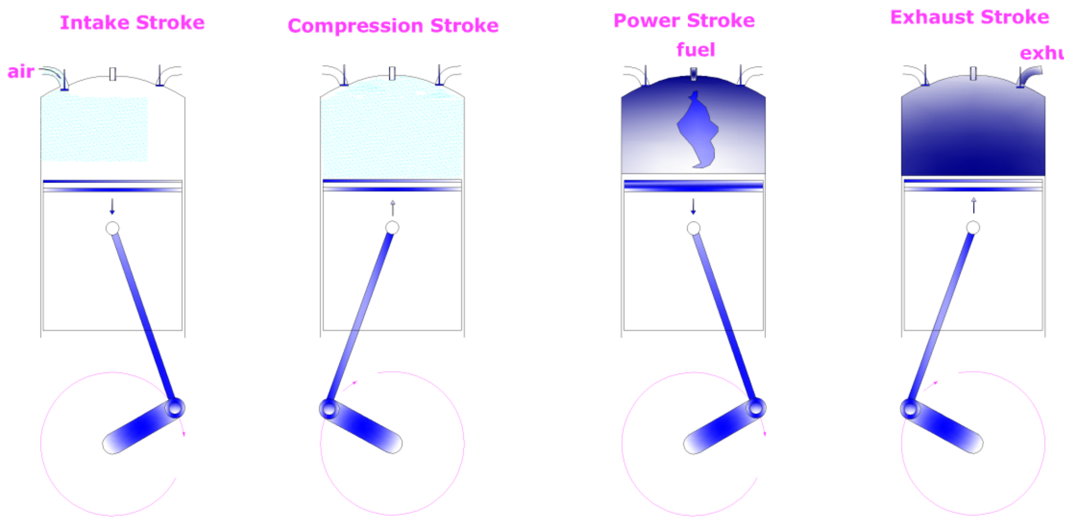


Figure 1.3 : Complete Ignition Cycle of SI Engine

at any instant. The relative piston position of four cylinders is shown in Figure 1.3.

Initially main focus of research in the field of SI engine was the improvement of "*Fuel Efficiency*". That was increased from 4% for engines built in early 1900 to 32% for engines built in 2000. With increasing number of vehicles, the problem of pollution on account of vehicles was aggravated so much that research stress is broadened to "*Reduced Exhaust Emission"*. To increase the fuel economy and control exhaust emission, the amount of fuel sprayed is controlled electronically. In Electronic Fuel Injection (EFI) vehicles the amount of fuel sprayed in engine is determined by the amount of air sucked in engine cylinder which is estimated by the position of throttle valve. The volume between the throttle plate and the intake valve of the cylinder is called intake manifold.

1.3 Major Components of an EFI SI Engine System

An EFI based SI engine is a complex system fabricated by an assembly of a large number of simpler components. Some of these components are listed below:

- Engine Cylinders
- Intake and Exhaust Manifolds
- Intake Ports and Exhaust Ports
- Fuel Injectors
- Igniter
- Electronic Control Unit

The optimal working of engine is ensured by a number of control loops present in engine. These control loops operate on the basis of information provided by a number of sensors present in engine systems. During suction stroke of an engine, the intake port opens and as the piston moves from TDC to BDC, air is sucked in the cylinder. The quantity of air sucked in engine cylinder is estimated using a Manifold Air Flow (MAF) sensor or Manifold Air Pressure (MAP) sensor installed in vehicles. The

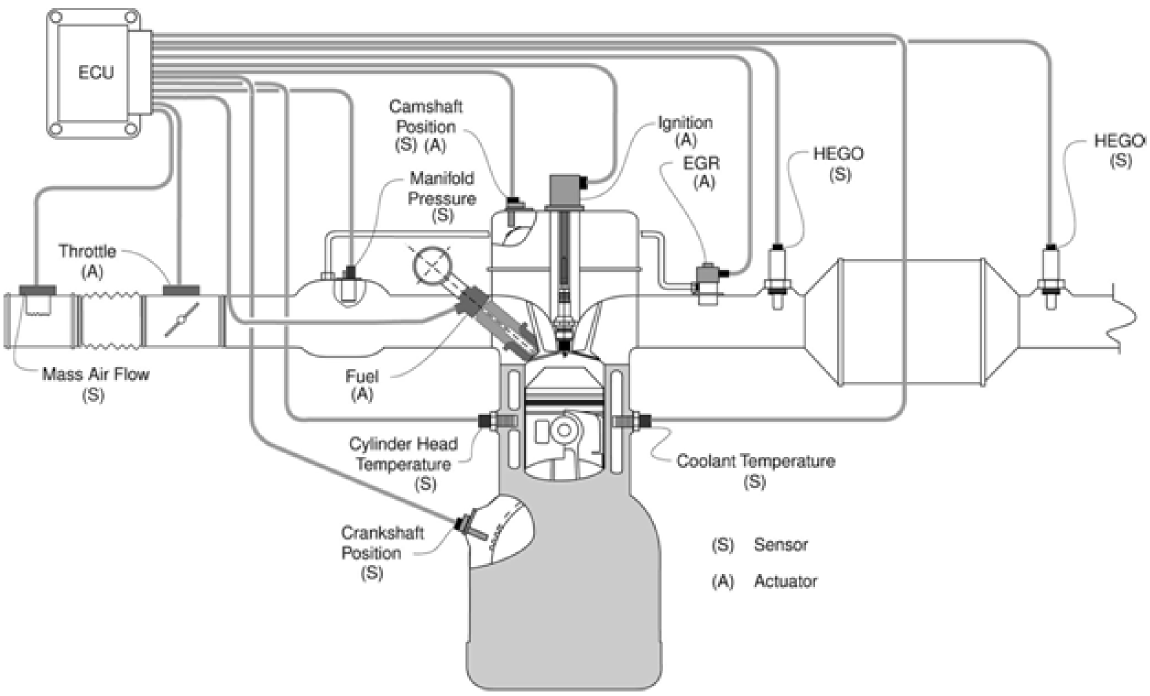


Figure 1.4 : Engine Cylinder connected with intake and exhaust manifold and sensors (*adopted from* **.** Cook J. A. etal, (2007, pp. 334 -335)

position of MAP sensor is shown in Figure 1.4. Crankshaft position sensor is also installed in all EFI vehicles. These sensors could be magnetic or optical sensors. A gear with known number of teeth is normally installed on crankshaft when magnetic sensors are used. When a tooth come close to the sensor, a magnetic coupling is established and a pulse is observed by the sensor. The gear assembly and position of crankshaft position sensor is shown in Figure 1.5. The internal structure of sensor and shape of its signal given by Stone R. *et al* (2002, pp: 135) are shown in Figure 1.5.

The information of number of teeth present on the gear assembly and the time taken by engine to traverse those teeth once can be used to estimate the engine speed. Figure 1.5 also indicates that most of the received pulses have same width but the width of one pulse is larger. This is due to a missing tooth on the gear assembly. This missing tooth is used to establish a reference point to identify the cylinder. In some engines, instead of a missing tooth the gear assembly contains a double tooth to establish the reference. In this case the observed signal would contain two narrow pulses instead of a single wide pulse. A throttle position sensor is installed in EFI engines. When the speed of vehicle is varied by changing the position of throttle, the sensor senses the new throttle position.

The estimate of throttle position, engine speed and air sucked in engine cylinder is passed on to the ECU. The ECU uses an internal lookup table to decide the amount of fuel to be sprayed in the cylinder. Under steady state conditions, the amount of fuel sprayed in engine cylinder is proportional to the amount of air sucked in it.

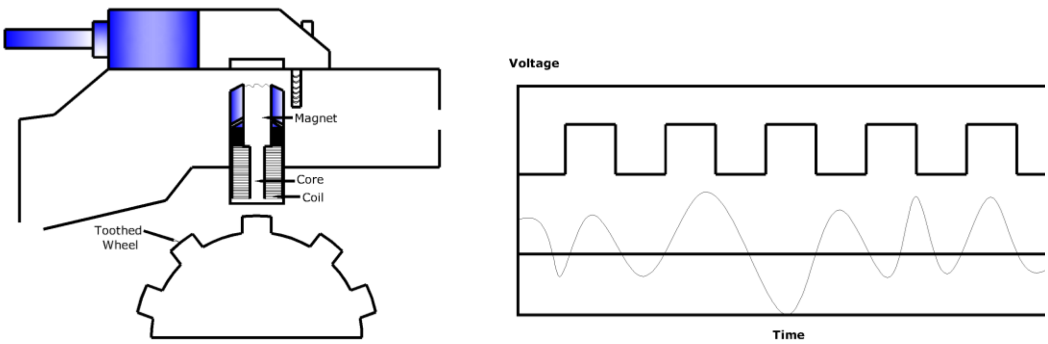


Figure 1.5 : (Left) Crankshaft Speed Sensor near gear teeth (Right) Sensor Signal

After the appropriate amount of fuel is sprayed in engine cylinder, the air fuel mixture is compressed. Under ideal conditions the compression is completed when the piston reaches the TDC and air fuel mixture would be ignited. In SI engine however it takes finite time for the burning of air fuel mixture and the pressure of burnt gasses is established after a finite delay with respect to the spark signal. A spark advance is maintained and spark signal is provided before the piston reaches TDC to ensure that piston would be at TDC when peak pressure of burnt gases is established in engine cylinder. The spark advance is also controlled electronically by ECU. On the basis of information from crankshaft position sensor, ECU decides about the ignition position.

After the completion of power stroke, the exhaust port opens and the burnt gases are pushed out of the cylinder. In the exhaust manifold, oxygen sensor is present to monitor the traces of remaining air and fuel in exhaust gases.

On account of analysis, the complex engine system is usually divided into six simpler subsystems:

- Air Intake Subsystem consisting of Throttle Valve and Intake Manifold
- Fuel Subsystem consisting of Injectors
- Engine Cylinders and Moving Assembly
- Engine Exhaust Subsystem
- Lubrication System
- Coolant System

Appropriate sensors and actuators required for correct working of each engine subsystem are present in engine. The actuator operation of EFI based SI engine is defined by the engine controller that provide control input on the basis of data provided by sensors. The introduction of faults in engine sensors, actuators or in engine systems may mislead the controller. Some fault scenario associated with engine system and the hazardous affects caused by those faults are given below:

1.4 Fault Scenarios in SI Engine

The major faults observed in an SI engine include:

- Ignition Fault
- Injection Faults

- Air leakage Faults

These faults can occur in only one of the engine cylinder or can occur in more than one cylinder simultaneously. To explain the effects of fault scenario, it is assumed that fault is present in only one engine cylinder and the other cylinders are providing sufficient power to ensure engine operation.

1.4.1 Ignition Fault

Ignition fault means that igniter signal is either not provided to a cylinder or the signal could not initiate a spark in cylinder when air fuel mixture was present in it. A direct consequence of problem is that ignition would not occur in cylinder and engine would not produce power in that cycle. The un-burnt fuel would be pushed to exhaust manifold in the next stroke of ignition cycle. The un-burnt fuel would be burned in catalytic converter of vehicle. An excessive heat generated in catalytic converter would damage it. If the fuel could not be burnt in catalytic converter, it would be exhausted to the atmosphere and cause pollution.

1.4.2 Injection Fault

Injection fault means either less/ no fuel was injected in cylinder (lean mixture formation) or excessive fuel is injected in cylinder (rich mixture formation). In case of formation of lean mixture engine would deliver less power. In the absence of fuel no ignition would occur and power would not be produced.

When a rich mixture is formed, engine would deliver normal power but some un-burnt fuel would escape to the exhaust system. This un-burnt fuel would be burnt in catalytic converter of exhaust circuit and gradually damaging it.

1.4.3 Air Leakage Fault

Air leakage fault has two different scenarios.

- Air leakage in manifold due to formation of some hole in manifold
- Air leakage from cylinders due to broken or lose piston rings

A negative air pressure is maintained in manifold. If a hole appears in manifold, more air would flow inside the manifold through hole and manifold pressure would increase. More air would be sucked in cylinder during suction stroke and engine would generate more power during the power stroke of the cylinder. This fault would affect the entire four cylinders equally in common manifold configuration.

When air is leaked from cylinders due to broken piston rings, the air-fuel mixture would leak out of the cylinder when the mixture is being compressed. This leakage would result in loss of power generation in ignition stroke of cylinder.

Misfire fault is one of the possible engine faults that result in loss of power. The faults like manifold air leakage fault, formation of rich mixture in cylinders etc. do not come in the domain of misfire fault. The misfire fault is therefore formally defined before defining the main problem.

1.5 Misfire Fault

The basic fault considered in this thesis is the ***Misfire Fault*** that represents the absence of formation of spark in engine cylinder as defined in the CARB report (2009, pp: 117). The main problems that result in engine misfire problem include:

1. Absence of formation of spark (fault in igniters or spark plugs)
2. Fault in air circuit (air leak from cylinders)
3. Fault in fuel circuit (Fuel pump or injectors inject less fuel in cylinder)
4. Any other fault that result in absence of formation of spark

Lee M. *et al* (2006, pp: 637-644) classified misfire into two major groups:

- Random Misfire
- Continuous Misfire

Random misfire occurs intermittently due to engine operation and road conditions and continuous misfire is due to fault in igniters, spark plugs, injectors and intake / exhaust problems.

Lee M. *et al* (2006, pp: 637-644) described the effectiveness of fault detection method as well as cost of fault diagnostic method as the two major points of consideration.

1.6 Statement of Problem

The proposed research work emphasizes on the "***Development of Novel Algorithm for Misfire Fault Detection in EFI based SI Engine using a novel Hybrid Model***". The fault isolation to attribute the fault to spark circuit, air circuit or fuel circuit is however not considered in this work and can be considered as the future extension to this work.

The basic assumption is that fault would be due to missing spark only and SI engine is operating in steady state conditions.

1.6.1 Hybrid Model

A Hybrid Dynamic System is characterized by the interaction of continuous and discrete inputs (Discrete Event) to the system. In these systems, the development of states variables of system for any given discrete event represents a *Mode*. The occurrence of discrete event therefore forces the system to change its mode and hence cause a jump in the values of state variables Messai N *et al* (2005, pp: 103-109). The basic challenge of working with hybrid systems is the measurement/ estimation of not only the continuous state of system but also identify the active mode Arogeti S. A. *et al* (2010, pp: 1452-1467). In case of hybrid systems a fault can be identified by the occurrence of an impermissible mode. The sequence of occurrence of events (modes) can therefore be analyzed to identify the fault. In case the mode switching events are well defined for a system FDI is easy Yu M. *et al* (2010, pp: 3000-3012). This is due to the fact that for deterministic and well known information of development of system modes, it is easy to estimate the system state variables and hence identify the faults. Arogeti S. A. *et al* (2010, pp: 1452-1467) mentioned that for accurate fault detection, fault parameter estimation requires information of system mode history. Messai N *et al* (2005, pp: 103-109) indicated that the identification of modes is a very hard problem in hybrid systems. A number of techniques like adaptive hybrid particle swarm optimization and bond graph approached are quoted in literature for the identification of mode switching and fault parameters estimation in hybrid system.

SI engine consists of four cylinders that can be modeled with continuous dynamics and a firing sequence of these four cylinders controlled by engine ECU represents a discrete dynamics. If firing event of SI engine is considered as a mode and crankshaft speed / acceleration are considered as the continuous state variables, then response of SI engine is defined by the interaction of both discrete firing event and the continuous dynamics of engine cylinders. The events of misfire can be considered as occurrence of an impermissible mode. Since firing sequence of an SI engine is deterministic, FDI for detecting misfire events could become simpler in the light of comments from Yu

M. *et al* (2010, pp: 3000-3012). The hybrid modeling is therefore a suitable option to represent SI engine for the detection of misfire fault. The problem would then be divided into the following steps:

- Mathematical modeling of engine cylinders (Modes)
- Mathematical modeling of ignition events (Events)

Using the observation of continuous data of engine speed, *Mode* of hybrid model can be estimated using stochastic techniques. The mode can finally be used to detect and identify the misfire fault of SI engine.

1.7 Proposed Solution to the Problem

A simplified hybrid model for SI engine would be developed on the basis of basic laws of physics. The input to the model would be considered to have some small randomly varying components like random fluctuations in fuel spray, burn time mentioned by Pulkrabek W. W. (1997, pp: 208-240) and amount of air sucked in engine etc. A stochastic analysis would also be carried out to study the effects of these random inputs in SI engine. The necessary data required for stochastic model would be obtained by studying the properties of hybrid model. Fault diagnostic algorithm would be developed using stochastic analysis of data. The scope of work of this thesis includes:

- Development of a novel Hybrid Model for SI engine and to study its properties.
- Stochastic analysis to study the effects of random variations of input on the output of model.
- Development of a novel fault diagnostic algorithm using Markov Chains
- Experimental validation Hybrid Model of SI Engine
- Experimental validation of Fault Diagnosis Algorithm
- Relative Operating Characteristic (ROC) analysis of fault diagnosis algorithm

1.7.1 Work Breakup

The work was initiated as a research oriented work. To validate the results of research, the arrangement of funds resulted in some additional commitments of some development work also. The study was therefore divided into two major areas i.e. research work and development work:

- The problem of research work of this study is defined in Section 1.6
- The development work is the Implementation of communication protocols of vehicles including ISO-9141-2 and CAN to develop a diagnostic toolkit for most vehicles being used in Pakistan.

1.8 Philosophy of Novel Fault Diagnostic Algorithm

Conventionally, in engineering systems, fault diagnosis is mostly carried out by forming a mathematical model and using the tools of control engineering for fault diagnosis. According to Biswas G. *et al* 2004 pp: 2159-2162, the community engaged in the development of fault diagnosis methods based on mathematical model and control engineering tools is named as Fault Diagnostics and Isolation (FDI) community. The community of computer science also needs fault detection in a number of their applications like in computer networks. This community is mainly using statistical techniques for fault detection and is named as DX community. Both the fault detection methods have their own advantages and suffer with some shortcomings also. An increasing trend is observed in the application of tools of DX community by FDI community and vice versa in their applications.

Jianhui L. *et al* (2009, pp:1-16) mentioned that neither model based nor data based methods can accurately solve the FDI problem and proposed an integrated approach based on parity equations, non-linear observer and State Vector Machine (SVM) for fault diagnosis in Antilock Braking System (ABS). The reference of simultaneous application of tools of both communities for fault diagnosis is however not widely available in literature.

The basic philosophy of this work is using simultaneous analysis of mathematical model and statistical properties of engine system to develop a computationally efficient misfire detection method. In this integration of model based and data based method; the properties of mathematical model provide heuristic guideline for the statistical method like data de-multiplexing, independence of states and enable us to prove that the selected random process is both Gaussian and Markov.

1.9 Major Issues in SI engine Fault Diagnosis

Fault is not the only source of error in a system but a completely healthy system may also exhibit error due to some unwanted inputs present in system. Some of these unwanted signals include input perturbation, disturbance and noise. Perturbation represents the slight change in input resulting in temporary departure of system from current state. Disturbance represents an unknown (and uncontrolled) input of limited frequency range acting on the system. Disturbance, being a physical signal is usually considered to be restricted to low frequency range. Noise is considered as unwanted input affecting the system response. The error on account of these signals result in some differences between actual system response and the model response.

SI engine with electronic fuel injection always operate in closed loop. The slight faults occurring in engine system are considered to be the disturbances by the engine controller. The controller generates the signals to counteract those disturbances. Wu, N. E. *et al* (1992, pp: 44-49) and Jacobson, C. A. (1991, pp: 22-29) discussed the role of controller in the design of diagnostic system in a closed loop. When an average behavior of engine is observed the effects of faults would not be observed due to the controller action. The engine *Mean Value Model* (MVM) can detect the fault only when the intensity of fault is increased to such an extent that engine ECU cannot handle it. The problem of detection of incipient fault in SI engine using MVM is therefore very difficult. The problem of detection of small abrupt faults using MVM is also a very challenging problem. The effects of disturbances like slight variation of air or fuel, statistical effects of forces acting on piston, formation of fuel film etc. from one ignition to other ignition are not important in MVM due to averaging nature of model. Any signal variation appearing on the sensor side is considered as noise and are frequently removed using additional filters in loop. The signal variation may however be actually due to some fault in engine system.

The main problem associated with the detection of misfire event is however the condition that the model must be capable of detecting and analyzing the events occurring within an ignition cycle. As misfire fault is associated with specific cylinders, the selection of hybrid model with independent subsystem corresponding to each cylinder would reduce the problem to study the response of individual subsystems.

1.10 Purpose of the research

The problem of fault detection in automotive systems is being studied since last two decades. The problem is still alive in the research community due to its complexity. The basic motivation for selection of proposed research has its roots in the lack of core knowledge of people affiliated with automotive industry in Pakistan. Although a number of vehicle manufacturing plants are present in Pakistan, the automotive research is a neglected area on part of both industry and academia. Also the automotive mechanics working in a large number of workshops in Pakistan lack even the basic knowledge of current trends of fault diagnosis in this field and still work on the basis of trial and error methods. Control and signal Processing Group (CASPR) was started in order to develop the research and development culture in Pakistan. The basic research interests of this group are modeling, control and diagnostics in automotive, fuel cell, aircraft, three tank system, robotics and radars.

The automotive group of CASPR was working on Mean Value Modeling of SI engine, parameter estimation and fault diagnosis in SI engine. Misfire fault was chosen due to its adverse affect on account of power and fuel losses in vehicle, environmental pollution and wear and tear of equipment due to jerky movement and burning of fuel in catalytic converter causing it to heat up and get damages. The effects of fault are so significant on account of the exhaust emission that after the resolutions about permissible exhaust emission were passed, the problem of detection of misfire fault was taken seriously and many vehicle manufacturing industries has funded the universities to identify its solution. The failure of igniter was included in OBD-II fault codes also. The spark plugs in modern vehicles are kept redundant to ensure that if one plug fails, the complete misfire would not occur. A simple solution that can easily be implemented in hardware is yet a problem and research community is still studying the problem. Most of the work is carried out either on the basis of mathematical models like MVM or DEM or on the basis of some signal based techniques. Since MVM is based on the average values of variables, it does not suit misfire detection applications. Also the complexity of DEM on ground of solving many nonlinear differential equations to define the response of a single ignition cycle make it inappropriate for integrating random input component with it to define the statistical properties of engine variables. A new simplified mathematical model for representation of power stroke of SI engine was therefore developed for study of statistical properties of engine variables and finally the results of statistical analysis were used in the development of a novel algorithm to detect the misfire fault. The chapters on modeling and fault diagnosis would provide a comprehensive survey of previous work in this area and development of proposed hybrid model.

A summary of works carried out during this study include:

- Development of mathematical model for representing SI engine
- Development of fault diagnosis algorithms
- Experimental validation of proposed model and fault diagnosis algorithm
- Development of strategy for implementation of algorithm on microcontroller

A part of research work presented in this thesis was published by author in his publications provided in section 1.13.

1.11 Applications of the research

For the application of research the existing fault diagnostic solutions available in local and international market were explored. A survey of automotive mechanics was conducted in different areas of Rawalpindi and Islamabad to explore their information regarding the engine operation and emerging fault diagnostic methods in the field of automotives. It was observed that most of the mechanics are either totally unaware of the existence of diagnostic capabilities of vehicles or they have just heard the news of some diagnostic devices. Even the workshops using the diagnostic devices bear only superficial information about the automotive fault diagnosis. The basic factors that restricted the flow of information to our local mechanics were explored and the two major factors were identified:

- Illiteracy of local mechanics who do not even know about the sensors
- The price of diagnostic equipment is beyond the reach of most of the mechanics

The information acquired during research is then utilized to develop low cost diagnostic equipment that the mechanics could purchase.

The work that has already been carried out on the topic as well as the continuation of presented work for future research and development would have a direct effect on the technical grooming of our local mechanics that represent a skilled worker community in Pakistan. The adaptation of improved fault detection method by the automotive mechanics would then have indirect effect on all the vehicle holders of Pakistan also. The work has already contributed to the technical uplift of a small local community of automotive mechanics.

1.12 Theoretical basis and Organization

The arrangement of thesis includes the definition of problem and identification of motivating factors behind the origin of research in the area is identified in chapter 1. A basic orientation of SI engine is also provided in first chapter. The research problem is formally formulated and the proposed solution is also identified in first chapter. Finally the major issues anticipated in research and applications of research are discussed. The next three chapters from Chapter 2 to Chapter 4 are dedicated to the literature survey. The actual work of author is presented in Chapter 5 to Chapter 8. A summary of work presented in different chapters is given below:

Chapter 2 provides the basic fault diagnosis terminology provided in literature. Chapter 3 provides a comprehensive survey of previous work in the area of misfire fault detection in SI engines. Different solutions of the proposed based on mathematical model as well as on signal and data based methods are discussed briefly. The proposed method is compared with some other methods found in literature on the basis of analytic simplicity and hardware requirements using literature survey. Chapter 4 gives some brief details of existing mathematical models of SI engine. In this regard, Mean Value Model (MVM), Discrete Event Model (DEM) and Data Based Models representing SI engine as a blackbox formed on the basis of Neural Networks.

Chapter 5 presents the Hybrid model after the development of inspiration for the development of model. The basic mathematical approaches found in literature for hybrid modeling are discussed. The approach of switched linear model is discussed in details and its properties are discussed. The statistically varying variables of model are identified and a stochastic analysis of crankshaft speed is performed and it was proved that crankshaft speed oscillations are Gaussian and Markov.

A brief introduction of Markov Chains is provided in Chapter 6 and the tool is then used to develop a fault diagnosis algorithm for the detection of misfire in SI engines.

To apply the algorithm, a set of states is first defined. These states are then used for residual generation for fault detection. It is established that the residuals could be analyzed using limiting probability of Markov chains to detect the misfire fault. It was finally concluded that the fault can also be identified using an intermediate result and a simple data based algorithm is established.

Chapter 7 provides the results and discussions on the work. In this chapter an overview of the data acquisition setup including engine and data acquisition cards is presented. The results of a set of four experiments are provided to illustrate the effectiveness of algorithm for the detection of misfire. After the establishment of affectivity of algorithm, the results were extended to explore the effectiveness of method for the detection of random misfire events and multiple misfire events. The analysis of true positive and false positive predictions is performed using Relative Operating Characteristic (ROC) analysis. The comparison of method with methods based on correlation analysis is provided at the end of chapter 7.

Chapter 8 is dedicated to the contributions and future work in which the basic contributions of this research work are discussed and applications of research are established.

1.13 List of Publications

- M. A. Rizvi, A. I. Bhatti, Q. R. Butt, “Hybrid Model of Gasoline Engine for Misfire Detection” Accpted for publication in “IEEE Transactions on Industrial Electronics” 2010.
- M. A. Rizvi, A. I. Bhatti, F. Malik, Q. R. Butt, “Hybrid Model Application for Fault Detection and Quality Assurance”, 16th International Conference on Soft Computing MENDEL 2010, Brno, Czech Republic, ISSN 1803 pp: 1803-3814.
- M. A. Rizvi, A. I. Bhatti, “Hybrid Model for Early Detection of Misfire Fault in SI Engines”, IEEE 13th International Multitopic Conference, 538315, pp: 1-6, Nov. 2009
- M. A. Rizvi, A. I. Bhatti, Q. R. Butt “Fault Detection in a class of Hybrid System”, International Conference on Emerging Technologies (ICET) 2009. ICET 2009. pp: 130-135, Oct. 2009

- M. A. Rizvi, A. I. Bhatti, Q. R. Butt "Misfire Detection in IC Engines using Finite State Automata", 15th International Conference on Soft Computing MENDEL 2009, Brno, Czech Republic, ISSN 1803 1803-3814, pp: 93-100, June 24-26, 2009
- Q. R. Butt, A. I. Bhatti, M. Iqbal, M. A. Rizvi, R. Mufti and I. H. Kazmi, "Estimation of Automotive Engine Parameters: Part I: Discharge Coefficient of Throttle Body", presented in IBCAST 2009, held in January 2009, Islamabad Pakistan.

1.14 Summary

The importance of Fault Diagnosis and Isolation techniques for SI engine is established and basic objectives of automotive research are defined in this chapter. After defining the effects of fault on the performance of SI engine the existing methods of fault indication in vehicles and their shortcomings are also explored. A broad classification of proposed fault diagnosis techniques is defined and major issues in fault diagnosis of SI engine are elaborated. With a brief overview of the problem in hand, the research problem is defined and the basic philosophy of the proposed solution is explained. Finally the importance of research is defined in the light of prevalent maintenance culture of vehicles in Pakistan and benefits of research are demonstrated. The organization of thesis is provided at the end of this chapter.

Chapter 2

FAULT DIAGNOSTIC TERMINOLOGY

The developments in the area of fault diagnosis and isolation were started as early as in 1970. The terminology used in this field was however inconsistent. In 1991, a steering committee called SAFEPROCESS (Fault Detection, Supervision and Safety for Technical Process) was formed within IFAC. This committee worked to develop a common terminology for the field under the supervision of R. Isermann. The proposed terminology is presented in Isermann R. *et al* (1992, pp. 709-719). The terminology presented in this thesis is adopted from the basic work of IFAC steering committee and the standard textbooks published on the subject in the next decades. This chapter presents the basic terminology of fault diagnosis, classification of faults along with its representation in system modeling and the basic structures of FDI algorithms.

2.1 Basic Terminology

2.1.1 Fault

Isermann, (2005, pp. 71-85) defined *"Fault as an un-permitted deviation of at least one characteristic property of a variable from an acceptable behavior* or Palade V. *et al* (2006, pp. 15-216) "*a deviation of system structure or the system parameters from the nominal situation*"

Israel Koren (2007, pp: 2-5) mentioned hardware defect in system components or a software bug in the firmware of system as two major sources of faults that would lead the system to some undesirable state. The presence of fault does not imply that the system has stopped working but it would result in system errors. System errors may be considered as manifestation of faults in a system. The fault can even be so insignificant that it cannot be detected easily yet it may be affecting the system performance. The representation of faults of different nature in a mathematical model is presented in section 2.3.

2.1.2 Failure

Mogens B et al (2006, pp: 8) expressed Failure as permanent interruption of a system's ability to perform a required function under specified operating conditions. A system may run for some time even after the manifestation of fault and before the occurrence of failure.

2.1.3 Fault Detection, Isolation and Identification (FDII)

Fault Detection means detection of fault in a system. ***Isolation*** means to identify the faulty component in a system. ***Identification*** means to identify the severity of fault in system. ***FDII*** represents the detection of fault, identification of faulty component and estimate the severity of fault in the system. For the case of detection of misfire fault in SI engine, the binary decision whether the misfire is present in any of the engine cylinder corresponds to fault detection, the identification of faulty cylinder corresponds to fault isolation and quantitative estimate of loss of affectivity of the defective cylinder corresponds to the fault identification.

2.1.4 Early Fault Detection

Early fault detection represents detection of fault at an early stage before the system degradation increased to an extent to cause any real damage in system performance. In fault diagnostic terminology it corresponds to the detection of incipient faults in system with an appropriate threshold defining the allowable system degradation.

2.1.5 Fault Prognosis

GeorgeVachtsevanos (2006, pp: 9-10) indicated that in fault prognosis future health of system and its components is predicted using currently available system data. Fault prognosis can be used to estimate the remaining useful life of a system. Fault prognosis is the key to the *Condition Based Maintenance* (CBM) in which maintenance plan is based on the observed machine health.

A survey of literature indicates that fault is classified on the basis of time or space.

2.2 Fault Classification Based on Time

On the basis of time response fault is classified into three basic categories as indicated in Ehsan S. (2009, pp:52) and Isermann (2005, pp. 71-85).

- Abrupt Faults
- Incipient Fault
- Intermittent Fault

2.2.1 Abrupt Fault

When the time between appearance and manifestation of fault is very small, the fault is said to be an abrupt fault. These faults are relatively easy to detect due to sudden large change in values of variables associated with fault. The value of variable can be compared with certain threshold value to detect the fault.

2.2.2 Incipient Fault

When a fault grows gradually in time, the fault would said to be an incipient fault. The effects of incipient faults become visible only after the magnitude of fault is increased above a certain threshold limit to cause some hazardous effects in system. To avoid the performance loss associated with faulty parameter during the time when fault is not manifested, it is necessary to detect the incipient faults well in time.

2.2.3 Intermittent Fault

These are the faults that appear and disappear repeatedly in time.
The time behavior of these three fault types is shown in Figure 2.1.

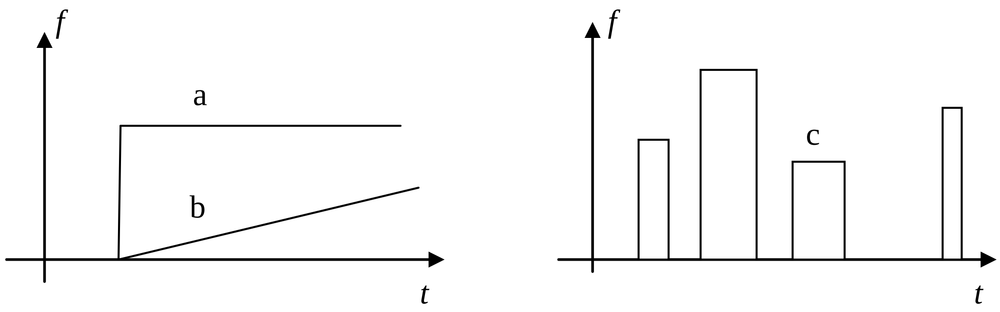


Figure 2.1: (Left) a. Abrupt Fault b. Incipient Fault (Right) Intermittent Fault

2.3 Fault Classification Based on Fault Location

Any plant can be divided into a number of sub-components. To monitor and control a plant sensors and actuators are also installed in the plant. Faults appearing in plant component are usually modeled as a change in plant parameters. The faults appearing in sensors and actuators are modeled as additive or multiplicative term in model. A classification of fault on the basis of fault location in a system is provided by Ehsan S. (2009, pp 22-24) and Ying H. *et al* (Technical Report 01-11-01)

- Sensor Faults
- Actuator Fault
- System Component Fault

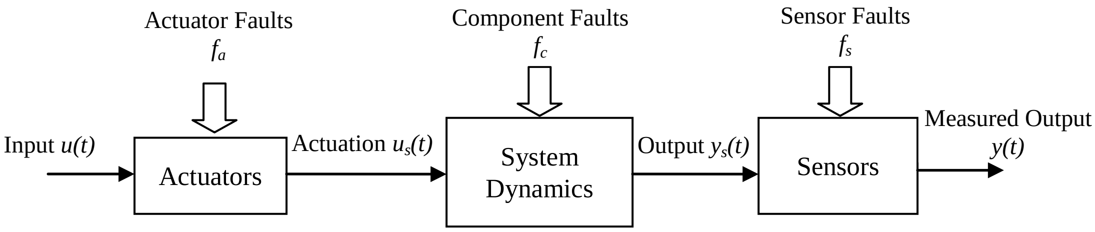


Figure 2.2: (Left) Actuator Fault (Middle) Component Fault (Right) Sensor Fault

The location of these faults in a system is shown in Figure 2.2, where f_a represents actuator fault, f_c represents component fault and f_s represents sensor fault. Hoffling *et al* (1996, pp: 1361-1369) and Isermann, (2005, pp: 71-85) described that faults that appear in technical systems can often be represented as additive or multiplicative faults with respect to the process model. Different faults associated with sensors, actuators and system and their representation in model are discussed below:

2.3.1 Sensor Fault

Sensor provide information about the internal states of a system to the controller and hence sensor fault will affect the decision making process of system which would then be reflected in degraded system response. The main faults associated with sensors are discussed by Ehsan S. (2009, pp 22-24) and are classified as (a) bias; (b) drift; (c) performance degradation (or loss of accuracy); (d) sensor freezing; and (e) calibration error. ***Bias*** represents a continuous offset in sensor output from the actual value. ***Drift*** represents that the sensor output changes with time and the error of sensor reading go on varying from the actual value. ***Loss of accuracy*** represents that sensor reading varies sufficiently from the actual value. ***Freezing*** represent that the sensor is showing a fixed value.

For model based fault diagnosis applications the fault would be modeled along with system and analysis would be performed on the integrated model of system and fault.

Since output is being measured by a sensor in a system so drift in a sensor will be represented as an additive fault on output side. If a feedback is present in system, then sensor fault will also result in some additive fault in input side of system. The structural representation of additive faults is shown in Figure 2.3.

Figure 2.3 : (Left) Additive Fault (Right) Multiplicative Fault

If the output of a system observed by a sensor is corrupted by a bias fault represented as an additive fault and the system controller passed the additive sensor fault to the actuator that also possesses a bias fault represented as an additive fault on input side, then the dynamics of system would be represented as:

$$\dot{x} = Ax(t) + Bu_R(t) \qquad \text{(Eq 2.1)}$$
$$y_R(t) = C\, x(t) + Du_R(t)$$

Where

$$u_R(t) = u(t) + f_a(t)$$
$$y(t) = y_R(t) + f_s(t)$$

Here, the signal $u_R(t)$ is the known input vector and the vector $y(t)$ is the measured output signal. These signals are corrupted by additive fault. The other sensor faults like drift and loss of accuracy are more difficult to model.

2.3.2 Actuator Fault

Actuators execute the controller command to drive the system. Actuator fault implies inability of execution of controller command. This may result in total loss of control. Actuator faults depend on the nature of actuator itself. The common actuator faults include Bias, Lock-in-place, floating, hardening or loss of effectiveness of actuator.

The actuator faults like bias may also be represented as an additive fault as can be seen from equation 2.1.

2.3.3 System Component Fault

Fault in a system component results in change in behavior of the system. This change in behavior may be represented as change in system structure, value of system parameter, location of system dynamic parameters like poles or zeros etc.

The change in values of system parameters is represented as multiplicative fault in system model. If *A*, *B*, *C* and *D* are the state space representations of system matrices, and due to some fault, the values of these parameters are changed by an amount ΔA_f, ΔB_f, ΔC_f and ΔD_f, then the resulting system would be given as:

$$\dot{x} = \left(A + \Delta A_f\right)x + \left(B + \Delta B_f\right)u \qquad \text{(Eq 2.2)}$$
$$y = \left(C + \Delta C_f\right)x + \left(D + \Delta D_f\right)u$$

For fault diagnosis applications based on frequency domain applications, a transfer function representation would be more appropriate. Considering *u(s)* and *f(s)* as the two inputs to the system, where *u(s)* is a regular input and *f(s)* is a fault input, the system output in the presence of additive fault can be defined using transfer functions of system input and fault inputs as:

$$G_u(s) = C(sI - A)^{-1}B + D \qquad \text{(Eq 2.3)}$$

and $$G_f(s) = C(sI - A)^{-1}R_1 + R_2$$

then

$$y(s) = G_u(s)u(s) + G_f(s)f(s)$$

The fault matrices R_1 and R_2 are usually assumed to be known [Chen J *et al* , (1999)]. Also the system parameters are considered for open loop i.e. the affect of controller is neglected in fault diagnosis scheme. In actual systems the accurate information of R_1 and R_2 is not available and also the controller cannot be removed from the loop.

Another problem in practical systems is that the input to the system is not known accurately. In the absence of actual input to the system, the reference command *r(t)* is normally used as input in FDI applications. The role of controller cannot be ignored. If a robust controller is being used for control applications, the minor faults would be treated as disturbance signals and the controller would generate the control signal to mask those faults.

A block diagram indicating system with reference input and a system having additive sensor fault at output, additive actuator fault at input and parametric faults in system parameters A, B and C is shown in Figure 2.4.

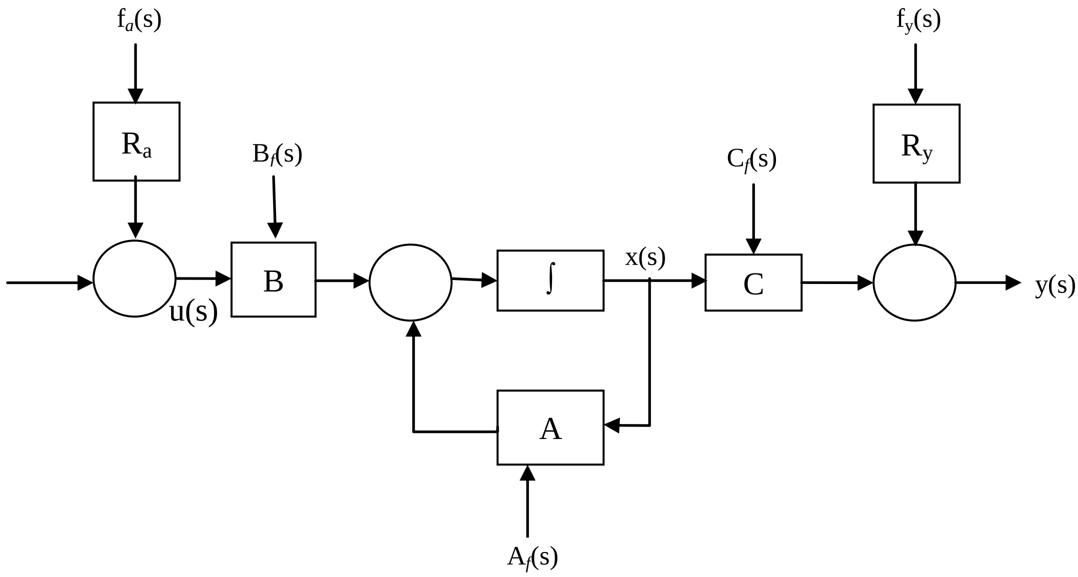


Figure 2.4 : System with Sensor, Actuator and System Faults

2.4 Fault Detectability

Nyberg M. (2006, pp. 1995-2000) defined Fault Detectability as the existence of a residual generator such that the transfer function from fault to residual is nonzero.

Consider that the system with system parameters *A*, *B*, *C* and *D* is excited by the input *u(t)* and an extraneous fault input *f(t)* is also acting on it. The output of system *y(t)* would be defined by (Eq 2.3). In these equations $G_u(s)$ and $G_f(s)$ represent transfer function for system inputs and fault inputs. The system faults are said to be ***detectable*** if G_f(s) is not zero. If the fault transfer function is zero the effect of fault would not be visible in system output.

2.5 Fault Diagnostic Methodology

A general survey of different fault diagnosis methods is provided by Patton R. J. *et al* (2000, pp: 298-311). The basic method is to compare the faulty system with a healthy system connected in parallel and operating under the same conditions, receiving the same input. Although this approach is feasible but it would not only result in additional cost but the approach is not always feasible due to system constraints like space limitations or process requirements. An alternate approach is to use *Analytic Redundancy* in system [Chow E.Y *et al* (1984, pp: 603-614)].

A system is said to be *Analytic Redundant* if there exists a functional relation between a measurement taken on the system and a variable representing that measurement in the diagnostic module of the system.

Most of the fault detection methods use some form of analytic redundancy relationships. The analytic redundancy can be expressed as dynamic mathematical model [Isermann R. (2005, pp: 71-85)], a data based model like an artificial neural network (ANN) [Ehsan S. (2009, pp: 22-24)] or the reference signals etc. Although the model of analytic redundancy varies, however the structure of all the fault diagnosis methods contains three major parts:

- Residual Generation
- Residual Evaluation
- Threshold Definition

Research work was carried out in all the three areas. A general block diagram representing basic structure of complete fault diagnosis methodology is shown in Figure 2.5. The *Process Model* block represents analytic redundancy. The input and output of actual system is provided as input to the Residual Generator. The residual generator calculates the difference of response of actual system and the model response. This difference acts as fault symptoms and is called *Residual*. The residual is then given as input to the residual evaluation block. The residual evaluations block analyzes the residuals in the light of some decision logic to detect the fault.

2.5.1 Residual Generation

Gertler J (2002, pp. 1-2) mentioned that the output of residual generator block is sensitive to faults. Under no fault condition the output of residual generator should ideally be zero. .In practice due to noise in measurement system and other factors, the residual is not zero. This makes the residual evaluation and robust fault detection a sufficiently challenging problem. The residual is required to have the following properties:

- Disturbance Decoupling
- Isolation Enhancement
- Resilience to noise

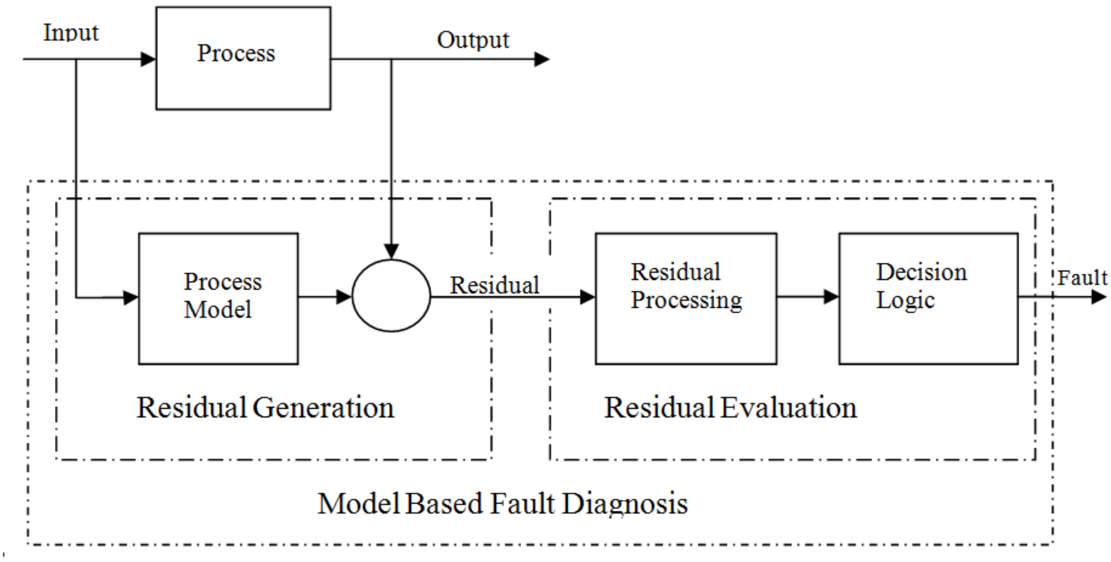


Figure 2.5 : Model Based Fault Diagnosis System

Disturbance decoupling means the generated residual should be insensitive to input disturbance but remains sensitive to faults. Isolation Enhancement means the residual should be generated in a way to provide features that can help in isolating the faults. Resilience to noise means residual should be robust enough to ensure the detection and isolation of faults even under noisy conditions. The output of residual generator does not depend on the operating point.

A simple residual generation method proposed by Gertler, J. et al (1988, pp.3-11) and Basseville (1988, pp. 309-326) is shown in Figure 2.6. In this approach $[H_u(s)\ \ H_y(s)]$

is made identical to the actual system, so that the signal z is the output of the simulated system. The difference between the output of the system and the simulator acts as residual. In this case however the stability of simulator cannot be guaranteed when the system in hand is unstable.

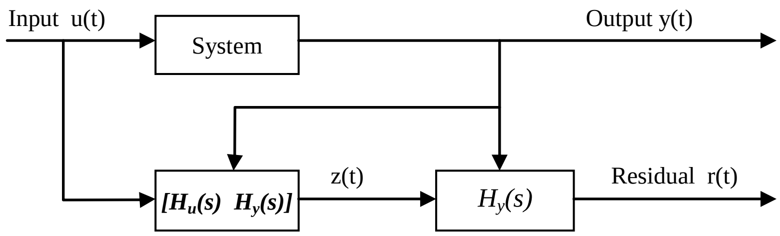


Figure 2.6 : Redundancy Signal Structure of a Residual Generator

Patton proposed an alternate structure for residual generation shown in Figure 2.7, which represents a generalized residual generator.

For a general residual structure shown in Figure 2.7, the residual output can be expressed mathematically as:

$$r(s) = \begin{bmatrix} H_u(s) & H_y(s) \end{bmatrix} \begin{bmatrix} u(s) \\ y(s) \end{bmatrix} = H_u(s)u(s) + H_y(s)y(s) \qquad \text{(Eq 2.4)}$$

$$\frac{r(s)}{u(s)} = H_u(s) + H_y(s)G(s)$$

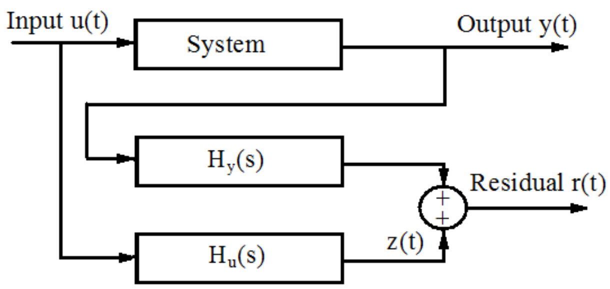


Figure 2.7 : General Structure of a Residual

Under no fault condition, the residual must be zero, and hence the condition for ideal residual generator is

$$0 = H_u(s) + H_y(s)G(s) \quad \text{(Eq 2.5)}$$

Patton, R.J. (1991, pp. 127-136) presented the problem of residual design as the selection of transfer matrices $H_u(s)$ and $H_y(s)$ that need to be realized using stable linear systems.

2.5.2 Nonlinear Residual Generation

The methods of residual generation discussed so far are valid for linear model of system. In practice, most of the systems show nonlinear response. For fault detection, a linear model is initially formed at the operating point and residual is generated using the methods discussed.

The method would however be applicable in a small range near the operating point. If the system operates in a wide dynamic range, the system could not be represented by the linearized model and the results may not be correct.

A number of alternate approaches for fault detection using nonlinear model are found in literature. These include introduction of neural networks for fault detection [Narendra, K.S. *et al* , 1990, pp: 4-27, Dexter, A.L. *et al* 1997, pp: 673-682] used Fuzzy logic integrated in model based FDI to overcome the problem of precision and accuracy of the models. The parameter estimation techniques using nonlinear model and sliding mode observer is providing more robust fault detection for non-linear systems. A new approach applicable to certain class of nonlinear system is to represent them as a hybrid model working as a piecewise linear switched systems Rizvi *et al* (2009, pp:130-135).

2.5.3 Residual Evaluation

The simplest technique for residual evaluation is based on defining some threshold limits for detecting the fault. For fault detection simple thresholds may be defined as:

$$y < y_{min} \ or \ y > y_{max} \Rightarrow Fault\ detected \quad \text{(Eq 2.6)}$$

$$y_{min} \le y \le y_{max} \Rightarrow No\ Fault\ detected$$

The residual signals are normally corrupted by the disturbances in the real system and uncertainties in system model. The robust residual evaluation techniques based on norms were therefore widely used for residual evaluation technique. The tools of robust control like $\mathcal{H}_2$ norm and $\mathcal{H}_\infty$ norms were hence used for residual evaluation.

The residual evaluation problem is defined as selecting suitable limits to encounter the effects of disturbances and model uncertainties. In this regard, the problem of threshold definition needs a clear definition maximum bound of exogenous inputs to system that affect the output. The bound of model uncertainties is also needed. If J represents some norm, J_{th} represents the largest possible value of J in the presence of disturbances d and model uncertainties Δ when no fault is present. The design of J_{th} that ensure the detection of fault in the presence of disturbances and model uncertainties can therefore be considered as an optimization problem:

$$J_{th} = \sup_{f=0,d,\Delta} J \qquad \text{(Eq 2.7)}$$

The problem of optimal selection of threshold values under the constraints of maximum disturbance or model uncertainty can be casted as an LMI problem to optimize certain norm. The description of threshold definition and inequalities used in the formulation of problem are discussed by Ding (2008, pp: 289-307)

Literature review indicates that threshold values based on techniques like fuzzy logic were also used instead of defining some crisp values.

2.5.4 Adaptive Threshold Method

For fault detection the residual would finally be compared with some threshold values developed on the basis of some criteria that are defined under same operating conditions and allowable disturbances. Hoffling and Isermann (1996. pp: 1361-1369), proposed that due to uncertainties in systems the residual always deviate from zero.

The basic reason behind the proposed argument is that it is very difficult to ensure that all the conditions of allowable disturbance are considered while calculating the residual. The comparison with the predefined thresholds may therefore result in some erroneous results.

In order to increase robustness in such cases, an adaptive threshold was proposed by Clark, R.N. *et al* (1989, pp. 231-233). The proposed methodology was later used in a number of applications. Pisu P, (2006, pp: 428-435) used this approach in diagnostic applications for "Steer by Wire", Witczak M. (2006, pp: 85-99) applied it in diagnosis

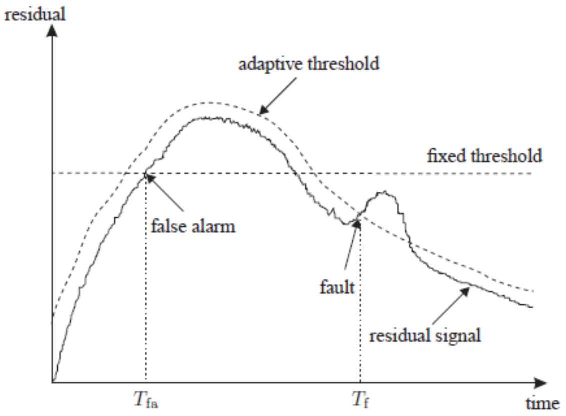


Figure 2.8 : Adaptive Threshold Selection

of control valve input. Figure 2.8 indicates the error on account of fixed threshold used for fault detection and how adaptive threshold selection can avoid this problem. Figure 2.9 shows a block diagram of adaptive threshold selection. [Emani-Naeini (1988, pp: 1106-1115)]

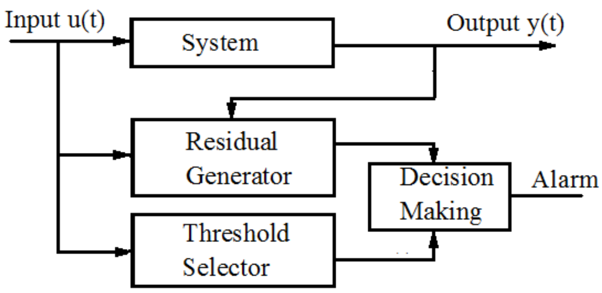


Figure 2.9 Adaptive Residual Selection

2.5.5 Residual Evaluation for Fault Isolation

When a number of faults need to be identified and isolated, it is difficult to find a single residual with distinct multiple threshold values corresponding to each fault. The problem of fault isolation is therefore a very challenging problem. A review of literature indicates that the problem is handled by defining multiple residuals each sensitive to one or multiple faults and defining a structure to isolate the fault. Ehsan S.

(2009, pp 31-32) described some residual evaluation structures for fault isolation that are described below:

2.5.5.1 Dedicated Observer Scheme (DOS)

In this structure a bank of residuals is defined with each residual sensitive to one particular fault and insensitive to all the other faults. The structure is proposed by Wünnenberg in 1990 (Ph.D. Thesis, Duisberg) and is shown in Figure 2.10. The fault isolation is therefore reduced to comparing each residual signal with the threshold value for the specific fault and identifying the fault using a Boolean decision table.

$$
\begin{aligned}
r_i(t) > T_i &\Rightarrow f_i(t) \neq 0 \\
r_i(t) \leq T_i &\Rightarrow f_i(t) = 0
\end{aligned}
\qquad \text{(Eq 2.8)}
$$

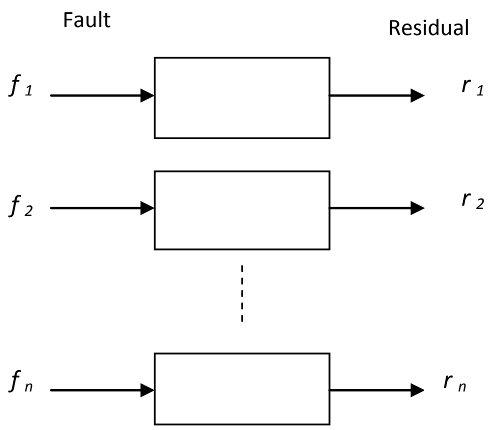


Figure 2.10 Structured Residual Set

The DOS scheme is good enough for the sensor faults. However, it has no robustness to unknown inputs like disturbance, uncertainty and noise.

2.5.5.2 Generalized Observer Scheme (GOS)

It is difficult to find a residual that is sensitive to a single fault and insensitive to all other faults. However one may develop residuals that are sensitive to many faults but remains insensitive to some particular fault. This forms the basics of generalized observer scheme, where a set of residuals are defined with every residual sensitive to all faults except one fault. The scheme is mathematically described by Ehsan S.

(2009, pp:35) and is given in (Eq 2.9). Fault isolation can then be performed by comparing the residual signals with the threshold values and identifying the fault using a fault table. A more detailed discussion about GOS scheme is given by Ehsan S. (2009, pp:35). The structure of GOS scheme for 3 possible faults is shown in Figure 2.11

$$\left.\begin{array}{r} r_i(t) \le T_i \\ r_j(t) > T_j \quad \forall j \in \{1, \dots, i-1, i+1, \dots, q\} \end{array}\right\} \Rightarrow f_i(t) \neq 0 \qquad \text{(Eq 2.9)}$$

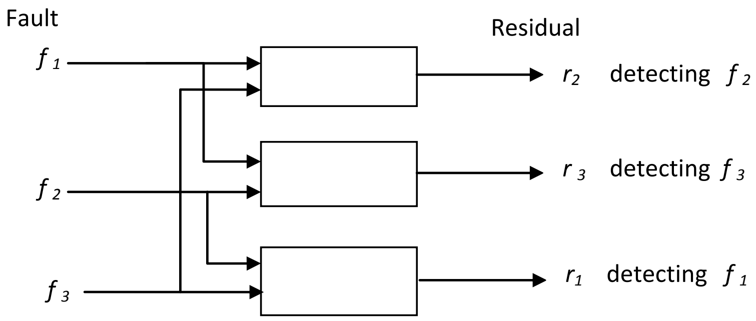


Figure 2.11 : Structured Residual Generator (Generalized Observer Scheme)

2.5.5.3 <u>Directional Residual Set Evaluation</u>

When multiple residuals are formed then residual can be represented in a vector space. Fault isolation can be performed by testing the direction of residual in the proposed vector space (Ehsan S. (2009, pp: 22-24)).

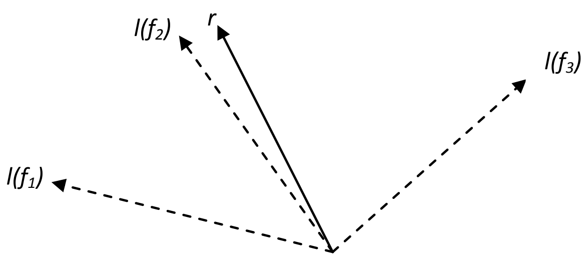


Figure 2.12 : Directional Residual Set

For fault detection some signature directions are identified corresponding to each fault and the direction of residual vector is tested to check how close the residual is to the defined signature directions. The Figure 2.12 shows the scheme indicating that f_2

is the most likely fault as the direction of residual *r* is very close to it. Other strategies of formation of structured residual using partial principal component mode is proposed by Gertler J *et al* (2004, pp: 1-3) and Gertler J et al DOI 10.1002/aic.00000.

2.5.6 Errors in a Fault Detection Method

The output of fault detection methods is a prediction of fault. The accuracy of prediction depends not only on the method but also on the conditions under which observations are being taken. Fahmida. M. C *et al* (2006, pp. 481-490) described that, under noisy conditions, a fault diagnosis algorithm may produce erroneous result. The detection algorithms may exhibit two types of errors:

- Misdetection
- False Alarm

Misdetection represents inability of algorithm to detect a fault, when fault is actually present in the system. ***False Alarm*** represents the weakness of algorithm to detect a fault condition when no fault is actually present in the system. The selection of stringent threshold conditions would result in increased rate of misdetection but reduced rate of false alarms and vice versa.

The problem of development of FDI algorithm can therefore be considered as an optimization problem to simultaneously reduce the misdetection and false alarm rate. The effectiveness of a fault diagnosis algorithm is therefore tested by the study of its prediction error i.e. False Alarm rate or Misdetection.

2.6 Qualitative Methods

A qualitative model describes the structured description of system on the basis of working principle. However instead of taking precise numerical values, states of system at important operating points are considered. Claudia M (1992, pp:1-14) applied this method for “Automated Rocket Engine Diagnosis”. Qualitative models using states of system can be applied for diagnosis of larger systems where it is difficult to develop a mathematical modeling of system. Mosterman P. J. *et al* (2000), indicated the superiority of qualitative method on the basis of mitigated complexity issues when compared with other numerical diagnosis approaches that need to study the transient behaviors in response to faults and convergence problems.

2.7 Fault Diagnosis in Hybrid Systems

As Hybrid Dynamic System is characterized by the interaction of continuous and discrete inputs (Discrete Event) to the system, the faults are not only caused by the change in system parameters of continuous system but also due to the disturbance/ faults in occurrence of discrete event.

Fault detection using Hybrid System is especially suitable for large, distributed and complex systems like process systems with fluid flowing in multiple tanks through multiple numbers of valves. The opening and closing of different valves act as system events and model of system under the conditions of specific valve openings result in development of different modes of system. Mander E. J (2002, pp: 235-240), developed a bond graph approach to model the system and use the dynamic characteristics of dependency relations between variables as a temporal causal graph. For each valve switching during a process cycle, the mode change occurs and system model switches. The parity relations for the system can be developed to detect and identify the faults using qualitative approach. This approach needs FDI design for each operating mode. Also only controlled mode changes using the assumption that only one tank can be draining and only one tank can be filling at a time. Yu M. *et al* (2010, pp: 3000-3012) mentioned that such method cannot work with unobservable events.

Arogeti S. A. *et al* (2010, pp: 1452-1467) indicated that inconsistency detected by residuals is not necessarily an indication of fault but may represent a mismatch between selected system mode and actual system. Accurate tracking of Mode is necessary for health monitoring of hybrid systems. Messai N *et al* (2005, pp: 103-109) mentioned that identification of current mode is a hard problem even when all nodes are known and discernible. The concept of analytic redundancy relations is extended on basis of global ARR (GARR) for application in hybrid systems. Arogeti S. A. *et al* (2010, pp: 1452-1467) used a Global ARR using model of hybrid bond graphs (HBG) to describe the behavior of hybrid system under all modes and finally used the rule based analysis of ARR to identify the mode.

The functionality of systems often prohibits the existence of system in some particle modes. Knowledge of valid system modes and prohibited system modes can be used to identify the system faults. The sequence of occurrence of modes also provides indication of faults. The identification of modes can be carried out using GARR based on system model, algorithms of pattern recognition or using data based techniques.

2.8 Summary

This chapter provided the basic terminology of the field of fault detection and Isolation. The chapter formally defined how fault is defined in an engineering process and identified different categories of fault on the basis of appearance of fault in time and location. It is also highlighted how different faults appearing in a system are represented mathematically for analysis of system.

A general perspective of a fault diagnostic algorithm is provided. Different components of a fault diagnosis algorithm like residual generation, residual analysis and threshold definition are identified. A literature survey is provided to identify the different structures being used in fault diagnostic algorithms.

The properties of residuals are defined it was discussed that most residuals are sensitive to multiple number of fault. A number of different schemes are identified that can process multiple residuals to correctly identify the faults. Literature survey is also provided to identify a problem in threshold generation that a fixed value of threshold cannot detect fault under all possible conditions of load variations and operating conditions and the concept of adaptive threshold condition is elaborated.

The error terminology of fault diagnostic algorithms like false alarms and mis-detection is discussed. The chapter was concluded with the discussion of diagnosis of faults in Hybrid Systems.

Chapter 3

MISFIRE FAULT DETECTION METHODS

This chapter presents a brief description of different methods found in literature that were applied for misfire fault detection in Spark Ignition (SI) engine.

The fault diagnostic algorithms were applied in all major engineering disciplines for the detection of faults in process equipments. Computer scientists used fault detection algorithms to achieve fault tolerance in software, data communication and hardware domain. Although the nature of problem was similar, the tools used by engineers and computer scientist were however quite different.

The literature classified the contributors of FDII as two independent fault diagnostic communities. The fault diagnostic methods are also classified on the basis of tools being used.

3.1 Fault Diagnostic Communities

Biswas G. *et al* (2004, pp: 2159-2162) described that development of Fault Diagnostic, Isolation and Identification (FDII) was carried out by two independent research streams. The domains of application of both of these communities were different and the tools used by them for fault detection were based on their domain of application. The literature referred these independent communities who developed the fault diagnosis algorithms in their domains as:

- FDI Community
- DX Community

3.1.1 FDI Community

The members of FDI community used the tools of control engineering for the detection and isolation of fault. Their methods are based on forming mathematical model of system under investigation. They used the methods of linear system theory like observers, fault detection filters, parity equations or parameter estimation to detect the faults. The tools of robust control like different norms, Linear Matrix

Inequalities (LMI) or non-linear methods like sliding mode technique [Jean J *et al*, 1991] were used for fault diagnosis.

3.1.2 DX Community

The members of DX community used the tools of computer science for the detection and isolation of faults. The DX community used *Data Based Techniques* and *Signal Based Techniques*. They used tools like statistical analysis, Bayesian networks, estimation theory, signal analysis techniques, wavelet analysis and neural network for fault detection.

Due to complexity of fault detection problem, the communal boundaries were gradually broken and both communities started applying the methods of other community in their applications to achieve simpler solutions. It is hence difficult to assign some fault detection methods to any one of these domain. As an example some model based method establishes their modeling using some previous knowledge of signal rather than developing the model on the basis of physical laws only. Bohn C (2005, pp: 239-244) used the information that crankshaft velocity exhibits some periodic oscillations and used the velocity harmonics to develop the model.

A coarse classification of FDI methods in the context of misfire detection problem of SI engine is presented in Figure 3.1.

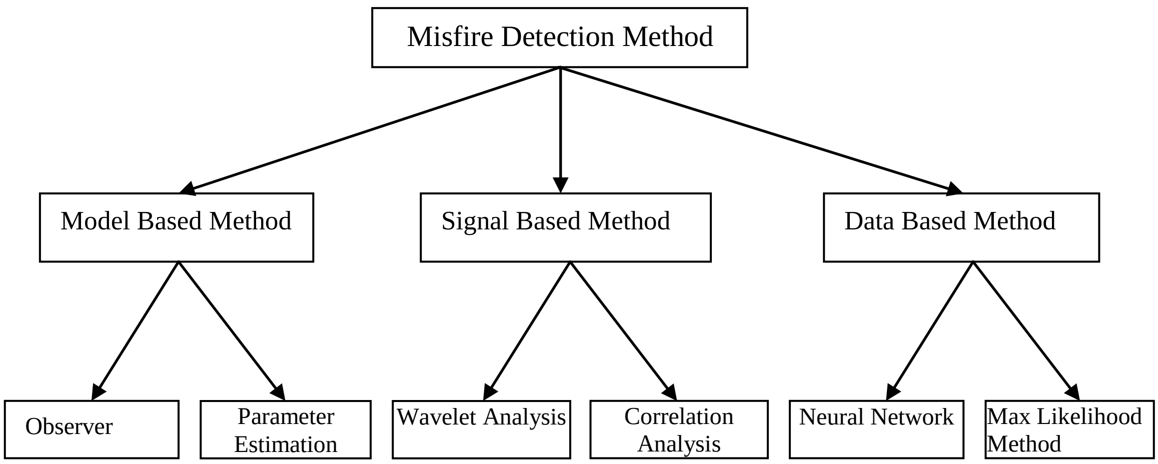


Figure 3.1 : Classification of Diagnostic Systems

A brief description of different methods used in misfire detection applications is provided in the coming sections of this chapter. The description of methods is provided in the context of their application to Misfire Fault Detection problem. A critical analysis of comparison of different misfire detection methods based on literature survey would be provided at the end of this chapter.

3.2 Model Based Methods

In *Model Based Fault Diagnosis*, an analytic model based on the basic laws of physics is first developed for the system under consideration. The analytic model acts as an analytic redundancy for detecting difference between the behaviors of the actual system from its desired behavior. This difference of behavior between the actual system and the predicted model behavior represents a *Residual*. Kiencke U *et al* (2005, pp-156) mentioned that performance of diagnosis can be increased by using model based diagnosis. Different methods of control engineering like state estimation, parameters estimations etc. can then be used to process the residual for necessary fault detection. Model-based fault diagnosis was first proposed in the early 1970s by Beard (1971, pp: 54-56) and Jones who developed Failure Detection Filters. Ding S. X, (2008, pp: 76-78) has given the details of some model based fault detection methods.

A number of references are found in literature for the misfire detection in SI engine using model based methods. These methods used the mathematical model of SI engine to represent engine torque, speed or acceleration as model output and input air, cylinder pressure etc. as model input. The behavior of SI engine is represented by a set of linear or nonlinear differential equations. Fault diagnosis is carried out by estimating the model outputs like torque and acceleration or the model inputs like in-cylinder pressure established during a power stroke.

The presented literature survey briefly describes the application of mathematical model for the detection of misfire fault. In this section, a brief description of mathematical model is also provided with the method.

3.2.1 Methods Based on Torque Modeling

The basic concept behind the misfire detection used in most of the model based method is that the effective torque produced by all the cylinders of an SI engine should be sufficiently balanced. Kiencke and Nelson (2005, pp: 183-184) defined residue as the relative error of the effective work given by balancing of torque as:

$$R_i = \frac{\frac{n}{4\pi}\int_{\theta_b}^{\theta_a} T_{ind}\, dt \; - \; \frac{1}{4\pi}\int_{t=0}^{4\pi} T_{ind}\, dt}{\frac{1}{4\pi}\int_{t=0}^{4\pi} T_{ind}\, dt} \quad \text{(Eq 3.1)}$$

$$where\ \theta_a = \theta_i + \frac{2\pi}{n} \quad and\ \ \theta_b = \theta_i - \frac{2\pi}{n}$$

In above expression θ_i represents crank angle on ith cylinder when piston is at the middle point between TDC and BDC. Under ideal conditions, all the cylinders are perfectly balanced and the residues are zero.

$$\sum_{i=1}^{n} R_i = 0 \quad \text{(Eq 3.2)}$$

When a specific cylinder generates less than average torque, the residue corresponding to that cylinder would become negative that indicate some fault associated with the cylinder.

Minghui and Moskwa (1994, pp: 2742-2747) estimated the torque generated by each cylinder stroke using mathematical model. For torque modeling a nonlinear varying inertia model for shaft given in Yaojung S. and Moskwa (1995, pp: 70-78) is considered. The detailed derivation of equation governing angular speed and torques are described in Yaojung S. and Moskwa (1995, pp: 70-78) and Kiencke and Nelson (2005, pp: 186). The governing equations are:

$$\dot{\omega} = \frac{1}{J(\theta)}\left(-\frac{1}{2}\frac{dJ(\theta)}{d\theta}\omega^2 + T_{ind} - T_{fric} - T_{load}\right) \quad \text{(Eq 3.3)}$$

$$T_{ind} = \sum_{i=1}^{n} T_{ind,i} = \sum_{i=1}^{n} P_i A_i L_{tor,i}$$

Where T_{ind} is the total indicated torque of all the four cylinders and could be estimated as the sum of torques due to individual cylinders. L_{tor} represents moment arm (perpendicular distance of axis of rotation from direction of application of force).

For torque estimation, a reference engine speed and load is provided. The nonlinearity of the rotating crankshaft dynamics is canceled by providing a combination of feedforward control term$\hat{T}_{load}(\omega_m)$ and $\left(\frac{1}{2}\frac{\partial J}{\partial\theta}\right)\omega_m^2$s, and feedback $\hat{T}_{fric}(\omega_m)+\hat{T}_{ind}$ and $\left(\frac{1}{2}\frac{\partial J}{\partial\theta}\right)\omega_m^2$ where ω_m is the reference speed. After decoupling of non-linearity through the mentioned terms, the equation 3.3 is reduced to:

$$\frac{\hat{\omega}}{\hat{T}_i} = \frac{1}{sJ(\theta)} \qquad \text{(Eq 3.4)}$$

The estimation of indicated torque is then reduced to a tracking problem of reference speed and load profile. A PI controller along with the mentioned feedforward and feedback terms to decouple nonlinearity is used to track the reference speed and load profile. The complete block diagram of the mentioned scheme is shown in Figure 3.2.

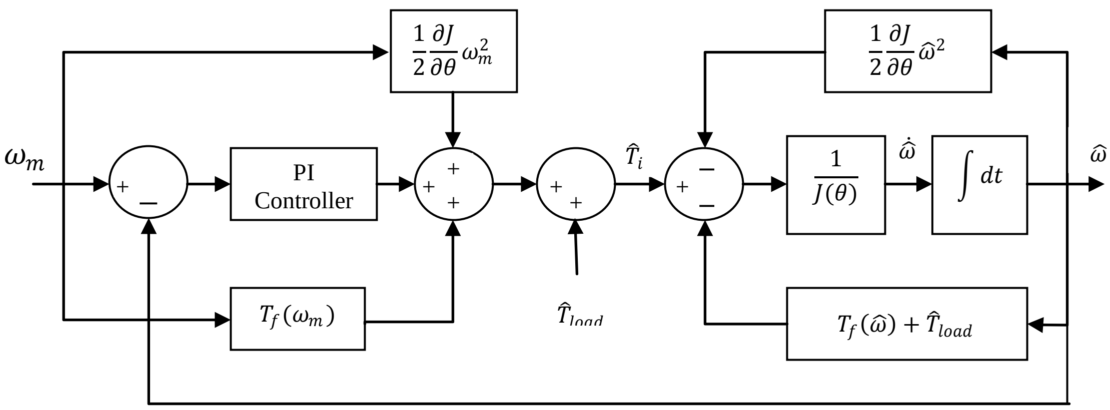


Figure 3.2 : Block diagram of PI feedforward indicated torque observer

Assuming an ideal decoupling, the characteristic equation of the closed loop system with PI controller to ensure the tracking is:

$$1+\left(k_p+\frac{k_i}{s}\right)\frac{1}{s\,J(\theta)}=0 \qquad \text{(Eq 3.5)}$$

$$s^2+\frac{k_p}{J(\theta)}s+\frac{k_i}{J(\theta)}=0$$

The gains of PI controller i.e. k_p *and* k_i were derived using pole placement. The input indicated torque is then estimated as the sum of nonlinear feedforward terms and PI control outputs as shown in Figure 3.2 and described by the equations:

$$\widehat{T}_i = \widehat{T}_f(\omega_m) + \widehat{T}_{load} + \frac{1}{2}\frac{\partial J}{\partial \theta}\omega_m^2 + k_p(\omega_m - \widehat{\omega}) + k_i \int (\omega_m - \widehat{\omega})\, dt \qquad \text{(Eq 3.6)}$$

Minghui and Moskwa (1994, pp: 2742-2747) also proposed the use of sliding mode controller in place of PI controller described earlier to ensure the tracking of reference speed and load torque. Walter A *et al* (2007) also used torque estimation for misfire detection.

Ball J. K et al (2000, pp: 1-24) also used estimated torque for misfire detection by measuring angular acceleration of engine block. Willsky A. S. (1976, pp: 601-611) mentioned the used of Kalman Filters for estimation using Stochastic models.

A misfire event in engine cylinder results in decrease in instantaneous torque. The direct measurement of torque however requires costly equipment. The methods developed for instantaneous torque estimation usually uses crankshaft speed. The accuracy of estimate are highly dependent on underlying mathematical model. The methods based on the estimation of torque using crankshaft speed are however more accurate under low speed and high load condition. Under high speed and low load condition, the inertia may mislead the algorithm causing false alarms. The basic advantage of using this approach is the availability of physical insight in fault detection method on account of underlying mathematical model.

3.2.2 Methods Based on Pressure Modeling

Under the assumptions of uniform distribution of temperature and pressure inside the cylinder after ignition, cylinder pressure dynamics can be expressed as a first order non-linear differential as described by Minghui and Moskwa (1994, pp. 2742-2747):

$$\dot{P}_{cyl\ i} = \frac{\gamma - 1}{V}\left[\dot{m}_{f\ burn}\, Q_{LHV} - \dot{Q}_{hi}\right] - \frac{\gamma}{V} P_{cyl\ i}\dot{V} \qquad \text{(Eq 3.7)}$$

In the above equation pressure is expressed in terms of rate of fuel burn and Lower Heating Value (LHV) of fuel. The equation can be solved analytically to get the solution:

$$P_{cyl\ i} = V^{-\gamma}(\gamma-1)V_{comb}^{\gamma-1}\left(m_{fburn}\,Q_{LHV} + Q_{hi}\right) + V^{-\gamma}P_{BDC}V_{BDC}^{\gamma} \qquad \text{(Eq 3.8)}$$

The solution of cylinder pressure clearly indicates two terms. The second term $V^{-\gamma}P_{BDC}V_{BDC}^{\gamma}$ represents polytropic compression of charge inside the cylinder and the first term accounts for the burning of fuel and heat transfer losses in the cylinder. The equation clearly shows that under misfire conditions, the first term would not be present. It is therefore expected that pressure difference between misfiring cylinder and healthy cylinder would be significant enough so that it can be used for the detection of misfire.

3.2.2.1 Methods Based on Pressure Estimation

Yaojung S. and Moskwa J. (1995, pp: 70-78) proposed a sliding mode observer to estimate the cylinder pressure due to combustion. The proposed pressure observer had two states i.e the instantaneous crankshaft angular speed and the cylinder pressure. The crankshaft instantaneous angular speed was measureable but the cylinder pressure of firing cylinder was assumed to be non-measurable. The nonlinear models of proposed states were:

$$\begin{aligned}\dot{\omega} &= \frac{1}{J(\theta)}\left(-\frac{1}{2}\frac{dJ(\theta)}{d\theta}\omega^2 + T_{ind} - T_{fric} - T_{load}\right) = f_1(\theta,\omega,P_1)\\ \dot{P_1} &= \frac{\gamma-1}{V}\left[\dot{Q_{ch}} - \dot{Q_{ht}}\right] - \frac{\gamma}{V}P_1\dot{V} = f_2(\theta,m_f,P_1)\end{aligned} \qquad \text{(Eq 3.9)}$$

where T_{ind} is the total indicated torque for all cylinders, V is the cylinder volume, P_1 represent pressure in the firing cylinder, T_{fric} is the mean friction torque, Q_{ch} represent the chemical energy of fuel and Q_{ht} is the heat transfer. Combustion heat transfer rate was estimated using lower heat value (LHV) and rate of burning of fuel mass. Fuel burning rate is calculated as the product of injected fuel mass and mass fraction burned. Injected fuel mass was estimated from the fuel flow rate and fuel injection duration and the mass fraction burned was estimated by a Wiebe function given as:

$$\frac{dQ_{ch}}{dt} = \frac{dm_b}{dt} Q_{LHV}$$
$$m_b = \left(1 - \mathrm{e}^{\left(-a\left(\frac{\theta-\theta_0}{\Delta\theta_b}\right)^{m+1}\right)}\right) m_f \qquad \text{(Eq 3.10)}$$

A sliding mode observer for pressure estimation was defined as:

$$\dot{\hat{\omega}} = \hat{f}_1 - \alpha_1 \widetilde{\omega} - K_\omega sat(\widetilde{\omega}/\eta)$$
$$\dot{\hat{P}}_1 = \hat{f}_2 - \alpha_2 \widetilde{\omega} - K_P sat(\widetilde{\omega}/\eta) \qquad \text{(Eq 3.11)}$$

where $sat(\widetilde{\omega}/\eta)$ is the saturation function. If Δf represents the modeling error between the observer model $\hat{f}$ and the true model f .The dynamics of estimation error was defined as:

$$\dot{\widetilde{\omega}} = \Delta f_1 - \alpha_1 \widetilde{\omega} - K_\omega sat(\widetilde{\omega})$$
$$\dot{\widetilde{P}} = \Delta f_2 - \alpha_2 \widetilde{\omega} - K_P sat(\widetilde{\omega}) \qquad \text{(Eq 3.12)}$$

A sliding surface was defined as the difference $(\widetilde{\omega})$ between measured and estimated engine speed. The linear terms with gains α_i was used as an aid to reach the sliding surface. Pressure estimates were obtained by appropriate design of gains K_ω and K_p. During sliding condition $s_1 = \widetilde{\omega} = 0$ and the system dynamics was reduced to:

$$\dot{\widetilde{P}} = -\frac{K_p}{K_\omega} \Delta f_1 + \Delta f_2 \qquad \text{(Eq 3.13)}$$

Using the appropriate choice of gains K_p *and* K_w under sliding conditions, the cylinder pressure was estimated. The results of pressure estimation using above method matched well with the experimental results given by author both under healthy conditions and faulty condition. The author successfully identified the misfire conditions by defining a threshold value of cylinder pressure to identify the faulty cylinder.

3.2.2.2 Methods Based on Pressure Centroid

Another model based method based on cylinder pressure model was proposed by Minghui and Moskwa (1994, pp. 2742-2747). In this method instead of defining a threshold point for cylinder pressure, the authors identified the location of pressure centroid in different cylinders of SI engine. Under misfire condition, pressure centroid would be close to zero when observed in any crank angle range symmetric about TDC (i.e. considering crank angle at TDC as reference zero and estimating pressure in an angular range $\mp\theta_0$). This is because under misfire condition no fuel would be burnt inside the cylinder and pressure profile would be more or less symmetric when observed in the mentioned crank angle domain. Under no misfire condition, the burnt fuel would produce high pressure inside the cylinder after ignition and the pressure profile in the range $[-\theta_0, 0]$ would be much lower than that in the range $[0, \theta_0]$. The centroid of pressure profile would therefore be shifted toward positive side. In this context pressure centroid is defined as:

$$C = \frac{\int P\theta d\theta}{\int P d\theta} \qquad \text{(Eq 3.14)}$$

In the above relation the numerator terms represent the first moment of pressure of each individual cylinders in the specified crank angle domain and denominator contain the average pressure over that crank angle domain. The sliding pressure observer used by Mingui *et al* (1994, pp. 2742-2747) for pressure estimation is:

$$\dot{\hat{\omega}} = \frac{1}{J(\theta)}\left[\sum_{i=1}^{n} 1000\hat{P}_{cyl\ i}\frac{dV_i}{d\theta} - \hat{T}_f - \hat{T}_{load} - \frac{1}{2}\frac{\partial J}{\partial \theta}\hat{\omega}^2\right] - \alpha_1\hat{\omega} - k_1 sgn(\tilde{\omega})$$

$$\dot{\hat{P}}_{cyl\ i} = \frac{\gamma-1}{V}\left[\dot{\hat{m}}_{fburn}\, Q_{LHV} + \dot{\hat{Q}}_{hi}\right] - \frac{\gamma\, \hat{P}_{cyl\ i}}{V}\dot{V} - \alpha_{cyl\ i}\tilde{\omega} - k_{cyl\ i}sgn(\tilde{\omega})$$
(Eq 3.15)

The effectiveness of the method was demonstrated by the authors using simulations and experimental results.

The in-cylinder pressure is the best method to identify the engine misfire condition. The sensors required for the measurement of in-cylinder pressure are however not only costly but their life is also limited due to the harsh environment in which they are installed. The estimation of torque also provide promising results but the torque estimate depends upon the model accuracy. The major advantage of this method is the physical reasoning provided by the method to explain the misfire conditions on the basis of underlying mathematical model. The major disadvantage is however the complexity of computations required for estimation of in-cylinder pressure. Yaojung S. and Moskwa J. (1995, pp: 70-78) on the basis of their work concluded that misfire detection based on cylinder pressure is more sensitive and convenient..

3.2.3 Methods Based on Acceleration Modeling

Rizzoni G. and Ribbin W. B (1989, pp: 423-436) stated that the information of both the average torque and time varying torque applied on the crankshaft of engine is available in the crankshaft acceleration signal.

$$\frac{\partial}{\partial t}\left(\frac{1}{2}J(\theta)\dot{\theta}^2\right) = \left(\sum T\right)\dot{\theta}$$

using $\ddot{\theta} = \frac{\partial\dot{\theta}}{\partial t} = \dot{\theta}\frac{\partial\dot{\theta}}{\partial\theta}$ and $\frac{\partial J}{\partial t} = \dot{\theta}\frac{\partial J}{\partial\theta}$ we get

$$\frac{\partial\dot{\theta}}{\partial\theta} + \frac{1}{2}\frac{1}{J}\frac{\partial J}{\partial\theta}\dot{\theta} = \frac{\sum T}{J\dot{\theta}}$$
(Eq 3.16)

Where θ is the crankshaft angle, $\dot{\theta}$ is angular speed of crankshaft, $J(\theta)$ is the moment of inertia as a function of angular position of crankshaft and $\sum T$ is the sum of all torque contributions. The combustion torque, load torque, friction and pumping losses are all lumped in the torque contributions.

The slight fluctuations of crankshaft angular speed results in the formation of periodic signal for angular speed. Bohn C. *et al* (2005, pp: 239-244) used the periodic nature of angular speed and used a Fourier expansion of the signal for modeling of speed and acceleration as:

$$\dot{\theta}(\theta) = a_0 + \sum_i a_i \sin\left(\frac{1}{2} i\theta + \alpha_i\right)$$

$$\frac{\partial \dot{\theta}(\theta)}{\partial \theta} = d(\theta) = \sum_i d_i(\theta) = \sum_i b_i \sin\left(\frac{1}{2} i\theta + \beta_i\right) \qquad \text{(Eq 3.17)}$$

Only a finite number of harmonics were considered and the individual harmonics $d_i(\theta)$ were modeled as a linear second order state space model, where $d_i(\theta)$ represents acceleration. A second order state space model (with angle θ as time variable) was established to model the individual harmonics.

$$\frac{\partial x_{d,i}}{\partial \theta} = A_{d,i} x_{d,i}$$

$$d_i = C_{d,i} x_{d,i}$$

$$where\ A_{d,i} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} i^2 & 0 \end{bmatrix} \quad \text{and} \quad C_{d,i} = [1 \quad 0] \qquad \text{(Eq 3.18)}$$

Representing the state equations of all the harmonics collectively in a matrix as:

$$\frac{\partial x_d}{\partial \theta} = A_d x_d \qquad \text{(Eq 3.19)}$$

$$d = C_d x_d$$

$$where$$

$$x_d = \begin{bmatrix} x_{d,1} & \dots & x_{d,L} \end{bmatrix} \quad A_d = \begin{bmatrix} A_{d,1} & & 0 \\ & \ddots & \\ 0 & & A_{d,L} \end{bmatrix} \quad and \quad C_d = [C_{d,1} \quad \cdots \quad C_{d,L}]$$

The state space model with crankshaft speed as a state variable was formed as:

$$x_1 = \dot{\theta}(\theta) \qquad d(\theta) = \frac{\partial x_1(\theta)}{\partial \theta} \tag{Eq 3.20}$$

The state space model of engine system was therefore represented as:

$$\frac{\partial x}{\partial \theta} = Ax \qquad where \quad x = [x_1 \quad x_d]' \qquad and \qquad A = \begin{bmatrix} 0 & C_d \\ 0 & A_d \end{bmatrix} \tag{Eq 3.21}$$

With N is the number of measurement sample per period, the system output was represented in discrete domain with index *k* that is an integral multiple of $\frac{4\pi}{N}$

$$y(k) = Cx(k) \qquad where \qquad C = [1 \quad 0 \quad \dots \quad 0] \tag{Eq 3.22}$$

The complete model representing the engine speed was therefore expressed as:

$$x(k+1) = G\,x(k) \tag{Eq 3.23}$$

$$y(k) = Cx(k) \qquad where \qquad C = [1 \quad 0 \quad \dots \quad 0]$$

Using the linearized model of engine angular speed in discrete angle domain, a state observer was designed as a Kalman filter. The equations of state estimators were:

$$\hat{x}(k|k-1) = G\hat{x}(k-1 \mid k-1) \tag{Eq 3.24}$$

$$\hat{y}(k) = C\hat{x}(k \mid k-1)$$

$$e(k) = y(k) - \hat{y}(k)$$

$$and \qquad \hat{x}(k \mid k) = \hat{x}(k \mid k-1) + K\,e(k)$$

where K is the stationary Kalman filter gain. The experimental results indicated in literature matched well with the results of estimator both under healthy and misfire conditions. For misfire detection instead of a direct comparison of results it was estimated that how much the estimator has to correct the estimated acceleration signal at each sampling instant.

The resulting acceleration correction signals were then plotted against the crank angle domain and the analysis of results indicated a negative peak value of correction signal when a cylinder misfire, followed by a positive peak when the next cylinder fires again.

The literature survey indicates some other model based techniques also. A cylinder pressure was estimated using sliding mode technique was proposed by Monterrubio J. M. *et al* (2007, pp: 620-624). A nonlinear engine model was formulated for estimation. The modeling strategy proposed by J.M. Monterrubio is however slightly different from that proposed by Moskwa. Another model based technique was proposed by Sood A. K. *et al* (1985, pp: 301-307) in which resultant torque was estimated by using a mathematical model based on the forces acting on piston. The difference between torques with and without fault was estimated and used for the detection of faults in engine.

The crankshaft position sensor is present in all EFI vehicles. The signal can be differentiated once to get an estimate of crankshaft speed. The mechanical construction of SI engine act as a low pass filter and the resulting crankshaft speed signal is sufficiently smooth. A slight variation in crankshaft speed however results in production of sufficiently large acceleration signal. The model for estimation of acceleration is most simple in all model based approaches and hence possible errors on account of model inaccuracies are very small for estimation of acceleration but due to second differentiation, noise enhancement would be significant.

3.3 Signal Based Method

Unlike Model Based Fault Detection methods where a system model is established on the basis of some basic physical principles to detect the faults, Signal Based Methods use the properties of the observed signal to detect fault. The basic assumption of signal based methods is that the observed signal would be deterministic with some known features. Some signal properties/ methods used for the development of fault detection algorithms are:

- Defining a signal space to represent signal using sinusoids, wavelets etc. and using the frequency domain analysis like FFT, wavelet transform or using some estimation techniques based on linear system theory.
- Signal modeling using Moving Average Model (MA), Auto Regressive Model (AR) or Autoregressive Moving Average Models (ARMA) etc.
- Defining a set of standard patterns of signal for faulty and healthy system and comparing the time domain signal with each element of the set to detect the fault.

Signal based methods use prior information of signal of healthy or faulty system without exploring the physical justification of signal on the basis of principles of physics. A brief description of some misfire detection algorithms based on signal analysis is provided in this section.

3.3.1 Methods Based on Moving Average (MA) Model

Lee A. *et al* (2003, pp: 3377-3381) mentioned that crankshaft speed is determined by the present speed, firing events (current) and manifold pressure and can be expressed as:

$$N(k) = f(N, P, I) \qquad \text{(Eq 3.25)}$$

Where N, P and I are crankshaft speed, average manifold absolute pressure and firing event signal vectors of dimension $n \times 1, m \times 1 \ and \ r \times 1$. The dimension of these vectors is defined by the contribution of past events on the current crankshaft speed. An inverse model for firing event is:

$$I(k) = g(N, P, I) \qquad \text{(Eq 3.26)}$$

The author acknowledged the difficulties in detailed modeling of nonlinear dynamic function G but claimed that for misfire detection the above inverse model can be simplified as:

$$I(k) = r(N)q(N,P) \qquad \text{(Eq 3.27)}$$

$$where \qquad N = [N(k), N(k-1), \dots, N(k-n)]$$

$$and \qquad P = [P(k), P(k-1), \dots, P(k-m)]$$

The function *r(N)* is termed as "*Engine Firing Event Estimator*" and the function *q(N,P)* is termed as the "*Load Compensator*". A misfire detection algorithm is proposed by Lee A. *et al* in the form of a moving average model as:

$$y(k) = r(N(k), N(k-1, \dots., N(k-m)) \qquad \text{(Eq 3.28)}$$

$$= b_0 N(k) + b_1 N(k-1) + \cdots + b_m N(k-m)$$

Where $b_i, i = 0, 1, 2, \dots \dots \dots, m$ are unknow model parameters to be estimated on the basis of test data from vehicle. The MA model given in equation 3.28 is written in state space form as:

$$y(k) = H(k)x(k) \qquad \text{(Eq 3.29)}$$

Where *x(k)* is (m+1) vector and *x(0) = b*, *y(k)* is the output and *H(k)* is a time varying measurement matrix.

Lee A. used a one step prediction Kalman filter for parameter estimation as:

$$\hat{x}(k+1) = \hat{x}(k) + K(k)[y(k) - H(k)\hat{x}(k)] \qquad \text{(Eq 3.30)}$$

$$= [I_{m+1} - K(k)H(k)]\hat{x}(k) + K(k)y(k)$$

$$with\ \hat{x}(0) = \hat{x}_0\ (arbitrary\ initial\ estimate)$$

$$and\ \ K(k) = \frac{\sum(k)H^T(k)}{H(k)\sum(k)H^T(k) + R(k)}$$

$$where\ \sum(k) is\ (m+1) \times (m+1)\ predictive\ error\ covariance\ matrix$$

$$\Sigma(k+1) = \Sigma(k) + Q(k) - \frac{\sum(k)H^T(k)\, H(k)\sum(k)}{H(k)\sum(k)H^T(k) + R(k)}$$

$$with\ \ \Sigma(0) = \Sigma_0\ \geq 0$$

Q(k) and *R(k)* are the known covariance matrix of the system noise *w(k)* and measurement noise *v(k)* respectively. The system equations are hence expressed as:

$$x(k+1) = F(k)x(k) + w(k) \qquad \text{(Eq 3.31)}$$

$$y(k) = H(k)x(k) + v(k)$$

The methodology was tested over a wide range of speed from 700 rpm to 6750 rpm under the loading conditions from no load to a wide open throttle and with no misfire to 100% random misfires. The results provided in literature indicates 100% success rate in the detection of misfire conditions with no false alarm.

3.3.2 Methods Based on Correlation Analysis

Rizzoni G. *et al* (1988, pp: 237-244) proposed a correlation based technique to detect and isolate the misfire fault. In this regard a number of data templates were formed by measuring the crankshaft speed fluctuations. It was assumed that each set consist of a periodic deterministic part with period six. For each fault condition, data was averaged over hundred cycles to approximate a deterministic template of fault representation. The templates were normalized for zero mean and unit power. A set of some templates of engine speed patterns developed under different misfire conditions and used by author for fault detection and isolation are provided in Figure 3.3.

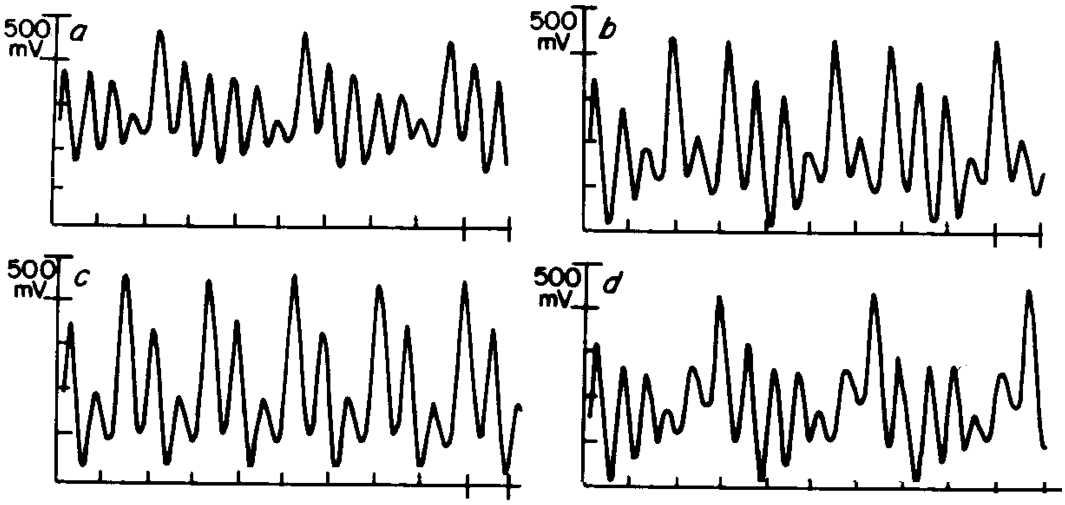


Figure 3.3 : Templates of engine speed patterns under different misfire conditions

Assuming a set of five templates, experimental data was then correlated with each of the five templates. This resulted in formation of a vector of dimension 5x1. The data set is then shifted by one step and a new correlation vector is formed. The process is repeated five times and a correlation matrix of dimension 5x6 is formed. The shifting is performed to discriminate the phase of the observed signal. The maximum element of the matrix provided the indication of fault.

3.3.3 Methods Based Wavelet Based Analysis

Matteo Montani *et al* (2006, pp: 144-148), proposed a wavelet based method to detect multiple misfires in SI engine using crankshaft angular speed signal. Speed was measured using magnetic sensor near the flywheel having 116 uniformly spaced teeth. The pulses of magnetic sensors were used to drive a counter. This signal was then analyzed to study the features needed to detect the misfire fault. Haar function was used as wavelet basis. The details of analysis filter blocks used for the detection of fault were however not mentioned in the paper.

The events of multiple misfire were detected by defining patterns corresponding to different misfire condition and finally analyzing the signal processed through filters with the known patterns.

3.3.4 Methods Based on Signal Behavior

Crossman J. A. *et al* (2003, pp: 1063-1075) analyzed the vehicle faults by relating them to signal behaviors. Crossman J. A. indicated that the vehicle responses can be expressed as a finite set of states and the signal exhibit different features in different states. Some of the features include abnormal magnitudes, rolling (smooth rise or fall), significant rise or fall, spikes, flat intervals or oscillations. Figure 3.4 indicates some of these signal patterns. Given a signal for larger time duration, the selection of appropriate time window for fault detection is called segmentation. Crossman J. A used a number of segments of interest and analyzed them for fault detection.

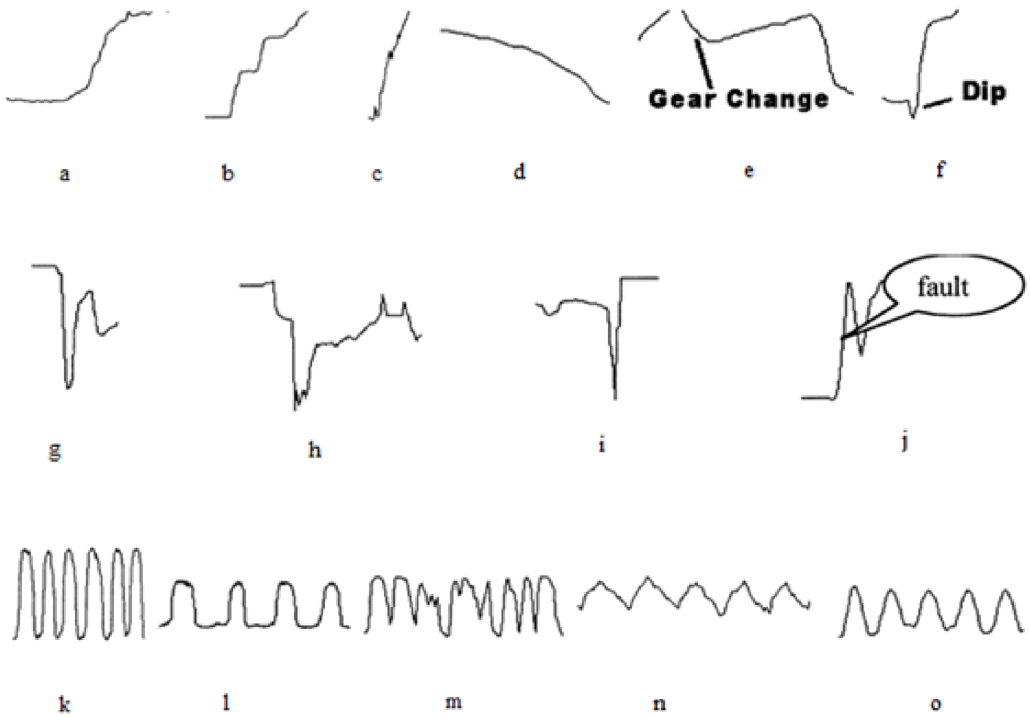


Figure 3.4 : Signal Features Associated with Systems

Murphy Y. L. (2003, pp-1076-1098) *et al* proposed a Distributed Diagnostic Agent System (DDAS) for automotive systems using Signal Diagnostic Agents (SDA). Each SDA diagnose one particular fault using single or a number of signals.

Parametric models including autoregressive model (AR) model were used by Rizzoni. The AR coefficients obtained using observation vector were used in testing the binary hypothesis of presence or absence of fault. The author described that the model order needed for correct identification of fault is 12.

A comparison of relative advantages and disadvantages of Signal Based Methods with Model Based Methods is provided in Section 3.5.2.

3.4 Data Based Method

Data based methods neither tries to neither find physical model of system under study nor assume a deterministic signal from the system. The output of the system is assumed to be a complete black box. Neural networks and neuro-fuzzy systems are the main representatives of such black box models (Holzmann H. *et al*, 1999, pp: 1014-1019). Other tools used in Data Based Methods include the basic tools of pattern classifications like Hidden Markov Model (HMM), Hybrid Bayesian Networks, Maximum Likelihood estimators etc. Details of some Data Based Techniques used for engine misfire fault detection are provided below:

3.4.1 Methods Based on Adaptive Classification

Feldkamp L. A *et al* (2000, pp:52-57) proposed an on-line learning system for input output data analysis and constructed a sequence of binary classification that work even in the absence of class information during training process. In the proposed method, the system is assumed to fall in one of the two mutually exclusive classes namely “Normal” and “Healthy” represented as 0 and 1 respectively. A set of neural networks were trained to model the input output behavior where each network modeled the system for one specific fault pattern. Fault state of the system was denoted by *s(k).* The fault states were defined by m bit patterns that represented current status of system and m-1 previous system status e.g. *m=3* means system can be in one of the 8 possible states say from 0 to 7. Since only current status of system can change, the system state can jump to only two possible next states from each state as shown in state transition diagram.

Assuming that current output of system is a function of input *u(k)* and state *s(k)* only.

$$y(k) = h(s(k), u(k)) \qquad \text{(Eq 3.32)}$$

Since the actual state *s(k)* is not known, it is termed as a hidden state and need to be estimated using the available data.

For p different fault patterns, p different networks were assumed. At each time step k, each of the p networks was provided with input vector $u(k)$ and the system response $\hat{y}_p(k)$ was estimated using Extended Kalman Filter (EKF). If σ_p^2 is the variance of error distribution of each network then the probability that a specific network represents the system is given by:

$$d_p(k) = \frac{1}{\sqrt{2\pi\sigma_p^2}}\, exp\left(-\frac{\left(y(k)-\hat{y}_p(k)\right)^2}{2\sigma_p^2}\right) \qquad \text{(Eq 3.33)}$$

A simple estimate of probability using above distribution may not ensure the evolution of states as described in state transition diagram shown in Figure 3.5.

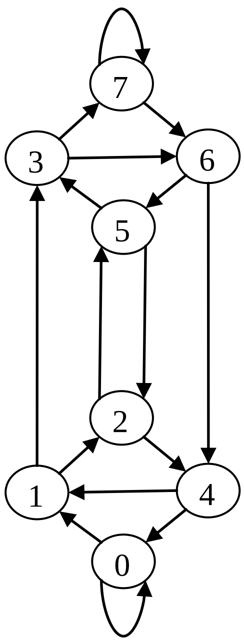


Figure 3.5 : State Transition Diagram

To infer the sequence of p, the forward backward procedure of HMM was used. The forward procedure estimated the probability of observing the sequence $y(1), y(2), \ldots, y(k)$ and state p at time k as:

$$\alpha_j(k) = \frac{d_i(k)\sum_i \alpha_i(k-1)a_{ij}}{\sum_l d_l(k)\sum_i \alpha_i(k-1)a_{il}} \quad where \quad \alpha_j(1) = \pi_j d_j(1), \quad j = 0,1,\ldots 2^m - 1 \qquad \text{(Eq 3.34)}$$

Where π_j is the prior probability of state j with l in the denominator run for all the states and k=2,3,…..T for a sequence of T steps. Here $\alpha_j(k)$ represent probability of occurrence of state p given the past events.

If $\beta_p(k)$ is computed in the backward part of the procedure where $\beta_p(k)$ is the probability of occurrence of sequence y*(k+1), y(k+2),….,y(T),* given that the starting state at time k is p. The update recursion for β_j is given by:

$$\beta_k(k' - 1) = \frac{\sum_i \beta_i(k')d_i(k')a_{ji}}{\sum_l \sum_i \beta_i(k')d_i(k')a_{li}} \qquad where \qquad \beta_j(T) = 1, \quad j = 0,1,\dots.2^m - 1$$

(Eq 3.35)

The probability $P_p(k)$ of being in state p at time k is estimated as:

$$P_p(k) = \frac{\alpha_p(k)\beta_p(k)}{\sum_{j=0}^{2m-1} \alpha_j(k)\beta_j(k)}$$

(Eq 3.36)

Network update for state of system is formed by scaling factor in proportion to the probabilities $P_p(k)$. The probability of fault is estimated by sum of probabilities of states with LSB equal to 1 (i.e. odd states)

$$prob_{fault}(k) = \sum_{p=Odd} P_p(k)$$

(Eq 3.37)

The variance of each state is updated every N_t steps ($N_t = 2000$) as:

$$\sigma_p^2 = \frac{\sum_k \left(y(k) - \hat{y}_p(k)\right)^2 P_p(k)}{\sum_k P_p(k)} \qquad and \qquad prob_{fault} = \frac{1}{N_t}\sum_k prob_{fault}(k)$$

(Eq 3.38)

The adaptive classification enabled the system to get trained with the incoming data and was used as a training algorithm for neural networks.

The proposed algorithm development was based on motivation of misfire detection problem to train a neural network and identify the fault conditions. Lee M. *et al* (2006, pp: 637-644) also proposed a misfire detection method based on neural network. The proposed method handled the problem in three stages: A data acquisition stage, signal preprocessing and feature extraction and pattern recognition for detecting misfire events using neural network. The preprocessing stage provided a correction for the effects of varying inertia on the signal. The measured engine speed was combined with nonlinear rotating dynamics to remove the affects of rotating inertia and generate the new synthetic variables of velocity and acceleration. These synthetic variables were then applied to a neural network for the detection of misfire fault. The network was trained using back-propagation method.

3.4.2 Methods Based on Likelihood estimation

Rizzoni G. (1987 pp: 450-457) proposed a data based technique for the detection of misfire fault in SI engines. The method is discussed in a more detailed manner as it provided an initial guidance for the research work carried out during this study.

The method is based on the concept of measurement of non-uniformity of engine torque and angular velocity. For the development of basic philosophy of method, a simple mathematical model was used with cylinder pressure as the input to engine. The forces generated by combustion process produce indicated torque defined by cylinder pressure P_i as:

$$T_i(t) = P_i(t).\,g(\theta) \qquad \text{(Eq 3.39)}$$

where $g(\theta)$ is a function of engine angular position and depend on engine geometry. $T_i\ and\ P_i$ are in general function of time. The geometry of engine imposes a periodicity with respect to crank angle on $T_i\ and\ P_i$. The study was carried out on steady state of a four stroke engine cycle for an angular range of $0 \le \theta \le 4\pi$.

Net torque acting on crankshaft is expressed as:

$$T_e(\theta) = T_i(\theta) + T_{fp}(\theta) + T_r(\theta) \qquad for \quad 0 \leq \theta \leq 4\pi \qquad \text{(Eq 3.40)}$$

where T_e is net engine torque, T_i is the indicated torque, T_{fp} is the torque lost in friction and pumping and T_r is the inertia torque contributed by the reciprocating assembly. Due to fluctuating nature of torque, the torque is completely specified by an AC component and a DC component. If T_{mean} is the mean engine torque then AC component of Torque would be estimated by subtracting the mean torque from the torque signal. Representing the torque fluctuations by τ, the AC component of torque can be expressed as:

$$\tau_e(\theta) = \tau_i(\theta) + \tau_{fp}(\theta) + \tau_r(\theta) \qquad for \quad 0 \leq \theta \leq 4\pi \qquad \text{(Eq 3.41)}$$

The relationship between torque and pressure inside the cylinder is expressed in terms of analogy between voltage in electrical circuits versus torque in mechanical systems and current in electrical circuits with angular velocity of rotating systems. A typical velocity waveform observed by measurement of crankshaft angular speed is provided and shown in Figure 3.6.

In his work Rizzoni mentioned that firing in each cylinder is associated with a maximum and a minimum in the waveform. He defined a vector of N extrema for kth ignition cycle is defined as:

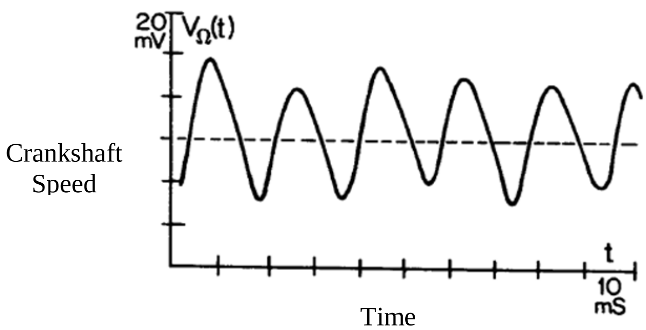


Figure 3.6 : Instantaneous Crankshaft Speed Signal as a function of time

$$\underline{\omega}(k)^{T} \triangleq [\omega^{1}(k) \quad \omega_{1}(k) \quad \omega^{2}(k) \quad \omega_{2}(k) \ldots \ldots \ldots \omega^{N}(k) \quad \omega_{N}(k)] \qquad \text{(Eq 3.42)}$$

where $\omega^{i}(k)$ represent i[th] maxima of k[th] ignition cycle and $\omega_{i}(k)$ represent i[th] minima of k[th] ignition cycle. He also defined a reference signal as:

$$\underline{\widehat{\omega}}^{T} \triangleq \|\underline{\omega}\|.[1 \quad -1 \quad 1 \quad -1 \ldots \ldots \ldots 1 \quad -1] \qquad \text{(Eq 3.43)}$$

with $\|.\|$ representing L_1 norm and $\underline{\widehat{\omega}}^{T}$ represents an ideally balance cylinder. The non-uniformity metric of actual cylinder from the ideally balanced cylinder was defined as the deviation of actual cylinder with ideally balanced cylinder as:

$$d(k) = \|\underline{\omega}(k) - \underline{\widehat{\omega}}(k)\|_{1} \qquad \text{(Eq 3.44)}$$

The proposed non-uniformity metric is a scalar quantity that would be zero for an engine with ideally balanced cylinders. For actual engines $d(k)$ is a random variable and $\{d(k),\ k = 1,2, \ldots \ldots .\}$ is a discrete random process. The value of non-uniformity metric was observed during N consecutive engine cycles and a histogram was plotted that approximately correspond to the probability density function (PDF) of random variable $d(k)$. The events $H_0\ and\ H_1$ were defined as:

$$\begin{aligned} \|H\|_{0} &= Engine\ operates\ normally \\ \|H\|_{1} &= Engine\ operates\ abnormally \end{aligned} \qquad \text{(Eq 3.45)}$$

Conditional probability density functions of $d(k)$ under the two events $H_0\ and\ H_1$ were defined as $P_d(d/H_0)$ and $P_d(d/H_1)$ respectively and are shown in Figure 3.7.

Knowledge of conditional probability density helped to define a threshold to decide between the hypothesis of engine operating under healthy condition or faulty condition. Having defined η as the threshold, a likelihood ration test (LRT) was defined by considering the function Λ, defined as:

$$\Lambda \triangleq \frac{P_{d/H_1}(d/H_1)}{P_{d/H_0}(d/H_0)} \qquad \text{(Eq 3.46)}$$

i.e. as a ratio of conditional probability density function of d under the hypothesis H_0 and H_1 respectively. The fault detection problem was then reduced to determine whether Λ is greater than η or not.

The Figure 3.7 also graphically indicates the interpretation of fault detection. The curve of probability density function is shifted with the occurrence of fault. The shifting of curve can be interpreted physically as the overall speed of engine is reduced with occurrence of fault and hence the distance of predefined reference points from new speed would become larger.

A detailed literature survey on the problem of misfire fault detection indicates the application of many more different techniques. Tinaut F. V. *et al* (2007, pp: 1521-1535) proposed misfire detection scheme based on engine energy model. Ball J. K (2000, SAE-2000-01-0560) *et al* used engine torque model for misfire detection. Different filtering and signal processing techniques were adopted for the improvement of misfire detection by Aono T. *et al* (2005, pp: 1218-1221), Stotsky A. A (2007, pp: 641-649), Naik S. (2004, pp: 181-198). State Observers for periodic signals was used for misfire detection by Bohn C *et al* (2007, pp: 641-649). Different methods based on algorithms of Pattern recognition like HMM, Neural networks and Support vector machines were developed by Wu Z. J. *et al* (1998), Lee M. (2006, pp: 637-644), Zhinong L. *et al* (2005, pp: 329-339), Devasenapati S. B *et al* (2010, pp: 25-29). Evolutionary computing control based on genetic algorithms was used for misfire detection by Kim D. *et al* (2007, pp: 3341-3355).

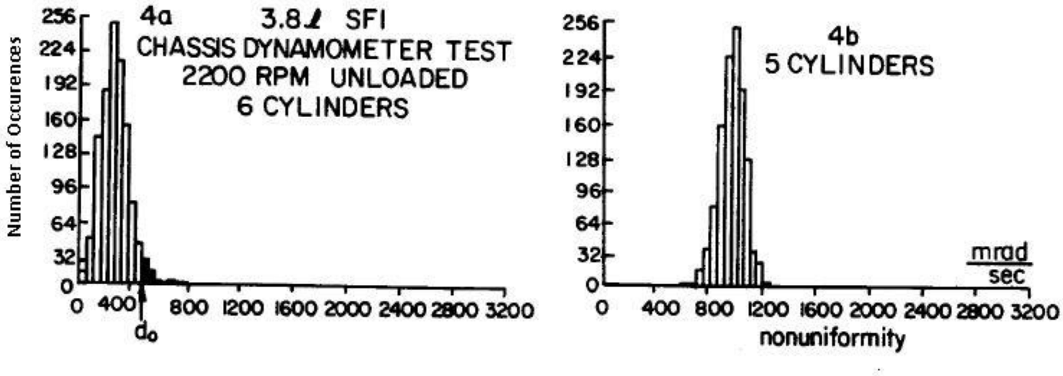


Figure 3.7 : Probability Densities of H_0 and H_1

3.5 Comparison of Different Misfire Detection Methods

In this section comparison of different methods of misfire fault detection in SI engine would be provided using literature survey. The comparison of methods would be carried out in the light of different classes of fault diagnostic algorithms presented in section 3.3 to 3.5. The merits and demerits of algorithms falling in specific classes would be considered. Another classification of the methods would be based on the requirement of sensors for the proposed method and a literature review would be provided to identify the suitable sensor for the problem in hand.

3.5.1 Merits and De-merits of Model Based Techniques

Sood A. K (1985, pp: 301-307) mentioned that model based technique has the advantage of physical reasoning to develop the classification rules for fault diagnosis. Using model based techniques, fault is detected either using state estimation [Minghui K. *et al* (1994, pp: 2742-2747), Yaojung S. *et al* (1995, pp: 70-78) etc.] or using parameter estimation [Sood A. K (1985, pp: 301-307)]. Literature review indicates that multiple parameters can be estimated using model based techniques [Butt. Q. R. *et al* (2008, pp: 3891-3898), Iqbal M. *et al* (2010)]. By carefully designing a model based techniques to ensure fault decoupling, these methods can be extended for the detection of multiple faults in systems. The robustness of fault isolation method would however be a real challenge in this scheme.

Ehsan S E. (2009, pp: 39) indicated that accuracy of mathematical model directly affect the diagnostic performance and reliability. He further mentioned that high fidelity mathematical models from physical principles can become very complicated, time consuming and even sometime unfeasible. Wong P. K. *et al* (2009, pp: 55-72) worked on automotive engine control problem and commented that SI engine is a complex multivariable nonlinear function that is very difficult to be estimated. The mathematical models based on physical principles are usually derived on the basis of many simplifying assumptions. These assumptions may lead to the structural change of model from the actual system resulting in significant modeling errors. As an example, Vemuri A. T. (2001, pp: 949-954) mentioned the availability of closed loop mathematical model as one of the key assumption that is not a valid assumption in practice. For an SI engine mathematical modeling on the basis of physical principles is carried out under open loop condition. An EFI automotive engine always operates in closed loop configuration with a highly robust controller present in the loop. The controller treats the engine faults as disturbances and would generate the control action to mask those faults. The detection of faults in closed loop when open loop system model is available is therefore a very challenging problem.

Even when the structure of model and actual system is same the accuracy of models depends on the correctness of parameter values [Holzmann H. *et al* (1999, pp: 1014-1019)]. Results of Butt Q. R. *et al* (2008, pp: 3891-3898) indicated that the model parameters are function of operating conditions of engine. The validity of model being used for detection is therefore itself questionable. Sood A. K (1985, pp: 301-307) also indicated this result and mentioned that this factor increases the complexity of fault diagnosis algorithms.

The methods of state estimation and parameter estimation can predict the status of current health of system however they are of limited help in prognosis applications.

In spite of the rapid development in the hardware when high speed processors and large memory blocks are available in small space, the fault diagnostic community still agreed that in a limited memory capacity of an ECU for the engine control applications, it is impossible to implement a comprehensive observer based residual generation scheme [Ding S. X *et al* (2009, pp:1-16), Jianhui L. *et al* (2009, pp: 1-16), Jianhui L. *et al* (2007, pp: 1163-1173)]

Another major issue in modeling perspective is the presence of discrete events in system that was completely ignored in mean value model of SI engine.

Most of the mentioned merits and demerits of model based techniques are quite general in nature and are applicable to most engineering systems. The most popular mathematical model of SI engine is referred as the Mean Value Model (MVM). Literature indicates that MVM is not suitable for diagnostic applications like misfire fault detection for which a Discrete Event Model (DEM) or a Cylinder-by-Cylinder model of SI engine would be required. [Karisson J (1998, pp: 1-8), Guzzella L. (2004, pp: 23)]. The computational complexity of these models is even higher.

3.5.2 Merits and De-merits of Signal and Data Based Techniques

The merits and demerits of Signal Based and Data Based Methods share sufficient similarity that they can be analyzed together.

The basic advantage in development of these techniques is that the effort required on account of model development would be saved. Some signals/ properties of signals may be identified on the basis of domain knowledge or expert opinions. Those signals may then be analyzed using fourier analysis Bohn C. *et al* (2005, pp: 239-244) or using algorithms designed on the basis of signal patterns.

Relatively simpler algorithms can be developed on the basis of Signal Based / Data Based Method that can be used for on-board diagnosis [Rizzoni G. (1987, pp: 450-457), Yu T. *et al* (1990, 53-57), Rizvi M. A. (2009, pp: 93-100)].

Data based techniques like Markov Chains can be used for predicting the future states of system given the present system state and hence data based methods are more suitable for prognosis applications as used by Morgan I. (2009, pp: 1774-1781). Similarly neural networks are considered as universal predictors and Manikandan V *et al* (2007, pp.82-91) has categorically indicated that a trained neural network can be used to classify a number of system faults.

The basic shortcoming of most data based algorithms that work on blackbox approach is:

- A mathematical/ physical link of method is difficult to comprehend and method cannot interpret the reason behind the fault.

- It is difficult to train the network/ collect classifiers under all possible fault scenarios.
- The fault detection algorithms may be misguided by the noisy data Angeli C (2004, pp: 12-30).

A comprehensive survey of different on line fault detection methods is provided by Angeli C (2004, pp: 12-30).

3.6 Comparison of Methods on the Basis of Sensors

A literature review of misfire fault detection indicates that a number of different sensors were used for fault detection applications. These include crankshaft speed sensor, cylinder pressure sensor, exhaust pressure sensor, oxygen sensor [Chung Y *et al* (1999, pp: 585-594)], vibration sensors [Villarino et al (2004, pp: 141-144)] etc. The cylinder pressure sensor is not only expensive but its life is also short. The cylinder pressure sensor, vibration sensor and exhaust pressure sensor are not normally present in production vehicles and need to be installed separately in engine. The crankshaft speed sensor is present in all production vehicles and was most widely used by the research community of misfire detection [Rezeka S. F. (1987, SAE Tech Paper 870546), Rizzoni G, (1988, pp: 237-244), Mauer G. F. (1990, pp: 221-226), Montani M. (2006, pp: 144-148), Kim D. E (2007, pp: 3341-3355)]. Lee M. *et al* (2006, pp: 637-644) mentioned that misfire fault detection using crankshaft speed sensor is difficult in small and medium sized engines because internal rotational inertia varies greatly with engine speed. The fault detection based on crankshaft speed fluctuations is really difficult and challenging under the low load conditions and high speed conditions. Lee M. *et al* (2006, pp: 637-644) also mentioned that the affects of variation of inertia can be removed from the signals by some preprocessing of signals.

Yaojung S. *et al* (1995, pp:70-78) worked on the misfire detection method using cylinder pressure sensors. They claimed that the method based on cylinder pressure is more sensitive and convenient. They however acknowledged that the methods based on crankshaft speed fluctuations are simpler, low cost and easy to implement. Yaojung also mentioned the problems of installation of cylinder transducers on vehicles. These problems include high cost of transducer hardware that prohibits their use in most production vehicles. The poor durability of these sensors is another factor that restricted their use in production vehicles in the past. The limitation of space in SI engine also restricts the installation of pressure transducer and associated equipment.

Merkisz J *et al* (2001, pp: 326-341) provided a brief survey of different methods of misfire detection along with their brief operational principle. He mentioned that exhaust gas pressure also exhibit oscillations and exhaust gas pressure falls rapidly due to the absence of combustion in cylinder. In case of misfire lack of pressure would be observed when exhaust port opens due to absence of high pressure in cylinders. The pressure sensor is however not present at exhaust in production vehicles. Merkisz also indicated some installation details about its installation in engine. He also discussed the experiments based on ionization current and optical methods. The work is concluded with the discussion of difference of different methods of misfire detection in terms of technical difficulties and cost. It was concluded that the methods based on crankshaft speed sensor are inexpensive and commercially viable but it is vulnerable to disturbances.

3.7 Current Status of Misfire Detection Problem

Murphey Y. L. *et al* (2003, pp:1076-1098) pointed out that automotive engineering diagnosis is still considered as the most challenging problem in engineering fault diagnostics. Although the comment is almost seven year old, yet a limited literature review provided in this thesis contain three papers accepted for publication in year 2010 on the problem of misfire fault in SI engine in three most reputed journals of engineering and automotive technology [Devasenapati S. B *et al* (2010, pp: 25-29), Malaczynski G. W. *et al* (2010, pp:1-11), Rizvi M. A. *et al* (2010,accepted for publication)]. This not only indicates that the problem is still under investigation by the research community but also indicates the importance and complexity of the problem that a satisfactory solution of problem is still being searched even when the problem has been studied for last two decades.

3.8 Application of Hybrid and Markov Models

Suchomski P. (2001, pp: 669-679) mentioned that in a hybrid system represented by the state space and a discrete set of system modes (where each mode correspond to a model), the mode jump define the changes in patterns of system dynamics. He suggested that finite state Markov Chains taking values in $M = \{1, \dots, M\}$ according to a proper transition probabilities is the most suitable representation of those jumps. Arogeti S. A *et al* (2010, pp: 1452-1466) mentioned that in hybrid system, the application of model based monitoring techniques is difficult due to unpredicted mode changes. Although the hybrid model proposed in this work has a deterministic switching of sub-systems, but under the influence of fault this switching pattern may change. Smith *et al* (2004, pp: 649-663) mentioned that there is limited work in utilizing Markov chains in fault diagnosis. Markov chains have the potential to predict the values of unmonitored dynamic variables and states. Sun S. *et al* (2004, pp: 437-441) used Markov chains for short term traffic flow forecasting where the volume of traffic flow in the next time interval has strong yet no deterministic relationship to the current state. Morgan I *et al* (2009, pp: 1774-1781) indicated that potential advantage of application of Markov Chains is its ability to predict. He applied it to predict the future concentration of elements in lubricant analysis of marine engine. The author mentioned that the area of application is critical in the sense that when ship is in sea, the maintenance is a big problem and early fault indication is therefore essential. The other area of applications of Hybrid Markov Models proposed by Yu X. G (2006, pp: 374-379) include the path prediction in mobile computing and wireless networks.

3.9 Emerging Trends of Fault Diagnosis Applications

Considering the difficulties in accurate modeling of complex physical systems and limitations of data based approaches, a new trend of using hybrid approach for fault detection is being adopted, that combines the computation of data based and model based techniques. Hybrid and integrated approaches are currently being studied for complex systems where formulation of an accurate and comprehensive mathematical model of system is difficult. Jianhui L. *et al* (2009, pp: 1-16) proposed an integrated model based and data driven diagnosis of automotive antilock braking system. Yan L. *et al* (2000, pp: 558-573) proposed method for learning the joint probability mass function (pmf) for discrete and mixed discrete/ continuous feature space. Ehsan S. *et al* (2009, pp: 16-17) proposed a hybrid approach to nonlinear fault diagnosis in which a priori information of system based on a mathematical model is used to create computationally intelligent techniques with adaptive and self learning capabilities.

A literature survey indicates that more powerful fault diagnosis algorithms are being developed by using the potential of data based fault diagnosis methods under the heuristic guideline of a mathematical model. The work of Lee M. *et al* (2006, pp: 637-644) used this approach by preprocessing the acquired data to remove the affects of varying inertia from the acquired data using an appropriate mathematical model.

3.10 Summary

This chapter provided a comprehensive survey of different approaches used for the detection of misfire fault. The methods of misfire detection were categorically classified on the basis of mathematical tools. The literature review indicated that most of the model based methods used estimation of torque or acceleration for the detection of misfire fault. The approximation of linear modeling was also mentioned where computational tools like Kalman filtering and observer design were used for analysis. The signal based method used the data of crankshaft speed and used correlation analysis, wavelet analysis, modeling using AR process etc to detect misfire fault. The methods of data based analysis were finally discussed in which the tools of pattern recognition like Bayesian decision, hidden Markov model and artificial neural network were predominantly used. After a comprehensive literature survey, the merits and de-merits of different methods were indicated.

The review of literature also indicated that most of the misfire detection methods are using crankshaft speed sensor to measure crankshaft speed and used it for fault detection. The major reasons observed in favor of its application were found to be its robustness, availability and cost.

Chapter 4

SPARK IGNITION ENGINE MODELS

A survey of literature on SI engine modeling indicates five different types of engine models to study the engine behavior.

- Mean Value Model (MVM)
- Discrete Event Model (DEM)
- Engine Kinematic Model
- Data Based Models
- Hybrid Engine Model

MVM is used most frequently by the research community working on SI engine for the development of control strategies and observer design. Data based models work on the basis of lookup tables formed and stored in engine ECU to control the engine operation. These lookup tables are formed predominantly on the basis of Mean Value Model. The hybrid engine model is not well mature model and author has presented a novel hybrid modeling approach in his research publications. In this chapter a brief description of some hybrid modeling approaches found in literature are discussed that are used to model SI engine. However the details of modeling work of author would be provided in the forth coming chapters.

4.1 Mean Value Engine Model

A comprehensive Mean Value Model (MVM) is defined by a set of three non linear differential equations defining the manifold pressure dynamics, fuel dynamics and rotational dynamics [Guzzella L. (2004, pp: 23), Weeks W. W (1995, pp: 1-15), Kim Y. W. (1998, pp: 84-99)], Hendrick E. *et al* [1990, SAE Technical Paper No. 900616]. The engine model with three states is highly complex and difficult to analyze mathematically. The computational complexity is simplified by ignoring the fuel dynamics with the assumption of stoichiometric air fuel ratio and using an engine model with two states only [Butt Q. R (2008, pp: 3891-3898)]

The basic manipulation variable of MVM is the engine throttle position and output variables are manifold pressure and engine speed. The modeling is based on average behavior of system states over multiple ignition cycles. The basic MVM is derived on the basis of basic principles of physics. A number of variants of MVM are found in literature depending on the basic set of assumptions taken while deriving the model. A brief description of basic set of manifold pressure dynamics and rotational dynamics of MVM is provided in this section.

Crankshaft rotational dynamics is expressed as:

$$I\dot{\omega}_e = T_{ind} - T_{load} - T_f \qquad \text{(Eq 4.1)}$$

Where T_{ind} represents indicated combustion torque, T_{load} represents load torque and T_f represents frictional and pumping torque, I is the moment of inertia of engine and ω_e represent engine speed. If H_u represents fuel energy constant, η_i is the thermal efficiency and $\dot{m}_f$ is the fuel mass flow rate in cylinder, then indicated-torque as given by Hendricks and Sorenson, 1990 is:

$$T_{ind} = \frac{H_u \eta_i \dot{m}_f}{\omega_e} \qquad \text{(Eq 4.2)}$$

The value of indicated torque varies with engine velocity. The frictional and pumping losses in engine is approximated as an empirical relation and was expressed as a polynomial in engine speed (Hendricks and Sorenson, 1990, Ganguli and Rajmani, 2004)

$$T_f = a_0\omega_e^2 + a_1\omega_e + a_2 + b_0\omega_e p_{man} + b_1 p_{man} \qquad \text{(Eq 4.3)}$$

Where a_0, a_1, a_2, b_0, b_1 are parameters dependent on the specific engine. The load torque is estimated using mean effective torque provided by the engine. After estimating the values of indicated combustion torque and frictional torques in equation (4.1), we come across the crankshaft speed dynamics.

If $\dot{m}_{ai}$ represents mass flow rate in the input manifold and $\dot{m}_{ao}$ represents mass flow rate out of the input manifold, then using law of conservation of mass on intake manifold, we get:

$$\dot{m}_{man} = \dot{m}_{ai} + \dot{m}_{ao} \qquad \text{(Eq 4.4)}$$

Pressure variations in intake manifold can be calculated using variations in mass flow rate using gas equation as:

$$\dot{p}_{man} = \frac{RT_{man}}{V_{man}} (\dot{m}_{ai} + \dot{m}_{ao}) \qquad \text{(Eq 4.5)}$$

The mass of air swept in the cylinder is defined by engine velocity that define air sucked in the cylinder. If V_{tdc} and V_{bdc} are the volume of air enclosed in cylinder when the piston is at TDC and BDC respectively, then the volume of air displaced/ sucked by piston would be given as:

$$V_d = V_{bdc} - V_{tdc} \qquad \text{(Eq 4.6)}$$

Mass of displaced air can be calculated using gas equation, however it depend on the breathing efficiency of engine cylinders called volumetric efficiency represented as η_{vol}. The equation of mass flow rate in the cylinders is therefore given as:

$$\dot{m}_{ao} = \eta_{vol} \frac{\omega_e}{4\pi} V_d \frac{p_{man}}{RT_{man}} \qquad \text{(Eq 4.7)}$$

For mass flow in the intake manifold, standard orifice equation for compressible fluid flow is applied [Cho and Hedrick, 1989] . A final equation presented by Ganguli and Rajmani, 2004 is:

$$\dot{m}_{ai} = MAX.TC(\alpha).PRI \qquad \text{(Eq 4.8)}$$

Where MAX is a constant depending on the size of throttle body and represents maximum possible intake airflow rate. $TC(\alpha)$ is the throttle chacteristic which is the projected area of flow as a function of throttle angle α and is modeled by Hendricks and Sorenson, 1990, Cho and Hedrick, 1989 as:

$$TC(\alpha) = 1 - \cos(\alpha + \alpha') \qquad \text{(Eq 4.9)}$$

Where α' is the minimum throttle angle seen be the engine when the throttle plate is closed against the throttle bore and PRI is “Pressure Ratio Influence Function”. In SI engines manifold pressure is always less then atmospheric pressure causing the air from atmosphere to flow to intake manifold through the orifice (Throttle valve).

$$PRI = \begin{cases} \sqrt{1 - \left(\dfrac{p_r - p_c}{1 - p_c}\right)^2} & p_r > p_c \qquad (choked) \\ 1 & p_r \leq p_c \qquad (sonic) \end{cases} \qquad \text{(Eq 4.10)}$$

where $p_r = \frac{p_{man}}{p_{amb}}$ and p_c is the critical pressure ratio that is approximately 0.5283. Therefore the manifold pressure equation becomes:

$$\dot{p}_{man} = \frac{RT_{man}}{V_{man}} \left\{ MAX.TC(\alpha).PRI\left(\frac{p_{man}}{p_{amb}}\right) - \frac{\omega_e}{4\pi} \frac{V_d \eta_{vol}}{RT_{man}} p_{man} \right\} \qquad \text{(Eq 4.11)}$$

4.1.1 Merits and De-merits of MVM

MVM is suitable for the applications of controller design. The model is however applied for fault diagnosis applications by estimating the model parameters and comparing it with actual values of those parameters.

In its simplified form, MVM is ignoring the fuel dnamics. Even in three state model the cylinder dynamics is completely ignored in MVM and mean effective torque is used. Many phenomenons normally visible in engine are not explained well using MVM. One such example is the slight speed fluctuations observed on engine crankshaft. Many parameters of MVM are unknown and are determined by forming an observer [Iqbal, Butt]. The results of parameter estimation indicates that parameter is not constant but is itself a function of engine operating conditions.

Guzzella L. (2004, pp: 23) described that in applications like misfire detection, where the event is not dependent on the average behavior of engine, MVM is not a suitable choice for fault analysis.

4.2 Discrete Event Model

A MVM is continuous and represent average behavior of engine. In a DEM all engine events correspond to the actual points in the cycle in which they occur. In MVM the independent variable is time and position is considered as the independent variable for a DEM. An IC engine is a discrete event system due to its reciprocating nature [Guezella]. The basic philosophy behind the DEM is to identify a specific crank-angle for each subsystems at which the actual boundary conditions (manifold pressure, air pressure, flow etc.) must be sampled.

To consider the basic philosophy of a DEM, consider torque produced by an engine is given by the relation:

$$T_e = \frac{H_l}{4\pi} m_{\varphi} . \eta_{eo}\left(m_{\varphi}, \omega_e\right) . \eta_{\lambda}(\lambda) . \eta_{\varsigma}(\varsigma) . \eta_{egr}\left(x_{egr}\right) \qquad \text{(Eq 4.12)}$$

Therefore engine torque can be influence by the following inputs:

- Mass of air in cylinder
- Fuel sprayed in cylinder
- Spark advance
- Exhaust gas recirculation rate

MVM considers the average values of variables but DEM considers the values of all parameters at given instant so considering the torque center as a reference point, the delay of center point of all the influencing variables are determined with respect to torque center to estimate the torque as:

$$T_e(t) = \frac{H_l}{4\pi} m_\varphi(t - \tau_{u_{inj} \to TC}).\eta_{eo}\left(m_\varphi\left(t - \tau_{u_{inj} \to TC}\right), \omega_e(t)\right).\eta_\lambda\left(\frac{m_\beta(t - \tau_{IC \to TC})}{m_\varphi\left(t - \tau_{U_{inj} \to TC}\right)}.\frac{1}{\sigma_0}\right).\eta_\varsigma\left(\varsigma(t - \tau_{U_{ign} \to TC})\right).\eta_{egr}\left(x_{egr}(t - \tau_{tra} - \tau_{IC-TC})\right) \quad \text{(Eq 4.13)}$$

DEM model expressed these delays in terms of crank angle. The main benefit of this change in variable is that in time domain time delays would be dependent on engine speed but in crank angle domain the time delays would be fairly constants on all engine speeds. The basic problem of discrete event model is to find the values of time delays. All time delays with respect to torque center are estimated for the calculation of torque. For torque estimation, the typical values of delays with respect to ignition are given by Guzzella (2004, pp. 141) and presented in Table 4.1.

TABLE 4.1 SPARK POSITION OF DIFFERENT STROKES OF IGNITION CYCLE

Engine Event	Position in˚ crankshaft angle after ignition
Intake Center (IC)	470˚ (110˚ after TDC)
Torque Center (TC)	80˚
Exhaust Center (EC)	250˚ (110˚ before TDC)
Update Ignition (U_{ign})	Defined in ECU $\phi_{update} = \phi_{offset} + \phi_{seg}.k$ k=0,1,2,3,….
Update Injection (U_{inj})	

For finding the torque at any time instant t, it is necessary to estimate the values of air intake, exhaust and injection at their respective timing and map their effect at time t. The computation of torque at any instant is therefore an iterative process. DEM is used to design controllers that require maximum bandwidth. Due to heavy computational load associated with the method, DEM was not used for the applications of fault diagnosis and isolation.

4.3 Kinematic model

Literature review indicates some engine models that consider the energy produced in engine cylinders as a result of ignition. The method is described briefly in this document as a part of proposed area of research of author is using some aspects of this model. A new simplified kinematic model derived by the author would be used in hybrid model proposed by author.

Maria P. F. *et al* (2003, pp: 1771-1776) used a model to represent the evolution of pressure in combustion chamber of a diesel engine. A similar model can be used to determine the force acting on piston. The model representing evolution of pressure in combustion chamber is based on first law of thermodynamics. When energy Q_i is added in i^{th} cylinder, then internal energy U of cylinder is added and some work W would then be performed using that energy. For notational simplicity, the index of cylinder would be neglected. The effect of change in internal energy is described mathematically as:

$$\frac{dU}{dt} = \frac{dQ}{dt} - p.\frac{dV}{dt} + \dot{m}_f h_{inj} \quad \text{(Eq 4.14)}$$

Where $\dot{m}_f$ represent fuel flow rate and h_{inj} represents the fuel enthalpy. By ideal gas laws, the internal energy of system is described as:

$$\frac{dU}{dt} = mc_v \frac{dT}{dt} \quad \text{(Eq 4.15)}$$

Assuming constant gas mass, state equations of ideal gas can be written as:

$$\frac{dT}{T} = \frac{dV}{V} + \frac{dp}{p} \qquad \text{(Eq 4.16)}$$

$$\Rightarrow \frac{1}{T}\frac{dT}{dt} = \frac{1}{V}\frac{dV}{dt} + \frac{1}{p}\frac{dp}{dt}$$

$$\Rightarrow mc_v\frac{dT}{dt} = c_v \frac{Tm}{V}\frac{dV}{dt} + c_v \frac{Tm}{p}\frac{dp}{dt}$$

$$\Rightarrow \frac{dU}{dt} = c_v \frac{Tm}{V}\frac{dV}{dt} + c_v \frac{Tm}{p}\frac{dp}{dt}$$

$$using\ pV = mRT, \qquad we\ get$$

$$\Rightarrow \frac{dU}{dt} = c_v \frac{p}{R}\frac{dV}{dt} + c_v \frac{V}{R}\frac{dp}{dt}$$

Substituting (Eq 4.15) and (Eq 4.16) in (Eq 4.14) and neglecting the heat component added due to enthalpy component $\dot{m}_f h_{inj}$, we get:

$$\frac{ldQ}{dt} = p\left(1 + \frac{c_v}{R}\right)\frac{dV}{dt} + V\frac{c_v}{R}\frac{dp}{dt} \qquad \text{(Eq 4.17)}$$

Using Mayer's relation $R = c_p - c_v$, we get

$$\frac{dp}{dt} = -\gamma\frac{p}{V}\frac{dV}{dt} + \frac{\gamma-1}{V}\frac{dQ}{dt} \qquad \text{(Eq 4.18)}$$

This is the basic model equation representing the cylinder pressure variations and the equations is same as (Eq 3.9) mentioned earlier in context with the work of Yaojung S. and Moskwa J. (1995, pp: 70-78) to estimate cylinder pressure variations during an ignition cycle. Having an estimation of cylinder pressure, it is possible to find the force acting on piston of cylinder as a function of crankshaft angular position as:

$$F(\theta) = p(\theta).A_p \quad \text{(Eq 4.19)}$$

where A_p is the piston area; The information of cylinder pressure and piston speed can be used to estimate the power being generated by the engine.

The forces acting on piston, causes it to move and this motion is then transmitted to flywheel through a crankshaft connected with piston. The mathematical model describing shaft angular velocity as a function of engine parameters and forces acting on piston due to burning gases is presented by Arun K. Sood. A simple line diagram of piston is shown in Figure 4.1. The force acting on piston can be divided into tangential and radial components as shown in Figure 4.1.

The piston position can be expressed as a function of angular position of shaft as:

$$x(\theta) = r + l - r\, cos\, \theta - l\, cos\, \phi \quad \text{(Eq 4.20)}$$

The velocity and acceleration of piston is given by:

$$v(\theta) = \frac{dx(\theta)}{dt}$$
$$a(\theta) = \frac{d^2}{dt^2}x(\theta) \quad \text{(Eq 4.21)}$$

From the Figure 4.1, using basic laws of trigonometry, it can be proved that:

$$l \sin\phi = r \sin\theta \quad \text{(Eq 4.22)}$$

Using (Eq 4.20) and (Eq 4.22), we get Eq 4.21 as:

$$a(\theta) = r\dot{\omega}f(\theta) + r\omega^2 g(\theta)$$

where $\omega = \dot{\theta}$ and (Eq 4.23)

$$f(\theta) = \sin\theta + \frac{r\sin 2\theta}{2l\left[1 - \frac{r^2}{l^2}sin^2\theta\right]^{\frac{1}{2}}}$$

$$g(\theta) = \cos\theta + \frac{r\cos 2\theta}{l\left[1 - \frac{r^2}{l^2}sin^2\theta\right]^{\frac{1}{2}}} + \frac{r^3\sin^2 2\theta}{4l^3\left[1 - \frac{r^2}{l^2}sin^2\theta\right]^{\frac{3}{2}}}$$

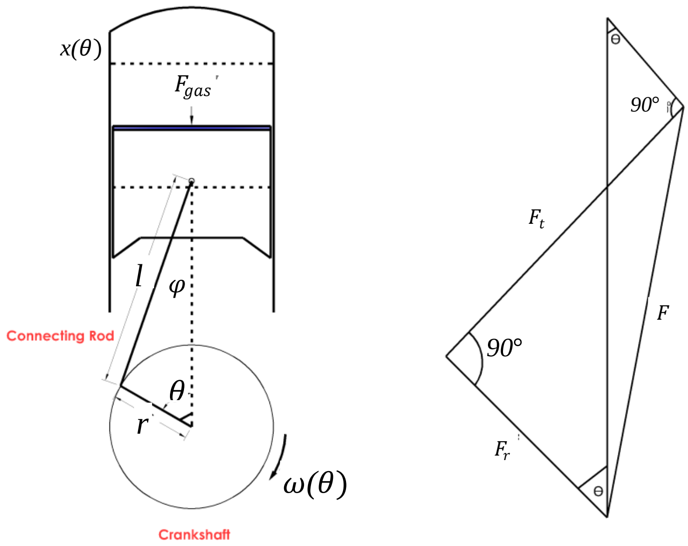


Figure 4.1: (a) Line Diagram of Crankshaft (b) Components of forces acting on crankshaft

The forces acting in the reciprocating parts are given as:

$$F_{IN} = ma(\theta) = F_{gas} - F_{fr} - F_c cos\phi \qquad \text{(Eq 4.24)}$$

Where

F_{gas} = Force on piston due to gas pressure

F_{fr} = Frictional force of reciprocating part

$F_c = Force\ transmitted\ through\ the\ connecting\ rod$

The force acting through the crankshaft can be divided into two components, one component in the direction tangential to the direction of motion of crankshaft and other in the direction radial to the crankshaft.

$$F_t = F_c \sin(\theta + \phi) \qquad \text{(Eq 4.25)}$$

$$F_t = \frac{1}{\left[1 - \frac{r^2}{l^2} sin^2\theta\right]^{\frac{1}{2}}} . \left[\left(F_{gas} - F_{fr}\right) - ma(\theta)\right]$$

From (Eq 4.25),

$$\sin(\theta + \phi) = \frac{r}{l} \sin\theta \cos\theta + \sin\theta \left[1 + \frac{r^2}{l^2}\ sin^2\theta\right]^{\frac{1}{2}} \qquad \text{(Eq 4.26)}$$

The net torque provided by the force F_c is given by:

- Torque due to load on engine
- Inertia of the rotating parts
- Torque due to friction in the rotating parts and accessories

Under no load conditions, the torque due to load is zero and torque would be given as:

$$I \frac{d\omega}{dt} = F_t r - T_{FR} \qquad \text{(Eq 4.27)}$$

Where *I* is the moment of inertia. Putting the value of F_t, we get:

(Eq 4.28)

$$I\frac{d\omega}{dt} = T_G - T_{fr} - T_{FR} - m\,G(\theta)a(\theta)$$

Where T_G and T_{fr} are the torques due to expansion of gas in cylinder and frictional torque. If

$$T_G = G(\theta)F_{gas} \qquad \text{(Eq 4.29)}$$

where:

$$G(\theta) = r\left[\sin\theta + \frac{r}{l}.\frac{\sin\theta\cos\theta}{\left[1 - \frac{r^2}{l^2}sin^2\theta\right]^{\frac{1}{2}}}\right]$$

Therefore we get

$$\frac{dw}{dt} = -\frac{G(\theta)g(\theta)}{\frac{I}{mr} + G(\theta)f(\theta)}\omega^2 + \frac{(T_G - T_F)}{mr\left[\frac{I}{mr} + G(\theta)f(\theta)\right]} \qquad \text{(Eq 4.30)}$$

Where T_F is the total frictional torque:

$$T_F = T_{fr} + T_{FR} \qquad \text{(Eq 4.31)}$$

Since

$$\frac{d\omega}{dt} = \frac{d\omega}{d\theta}\frac{d\theta}{dt} = \omega\dot{\omega} \qquad \text{(Eq 4.32)}$$

Therefore

$$\dot{\omega} = b(\theta)\omega + \frac{T_G - T_F}{d(\theta)}\omega^{-1}$$

(Eq 4.33)

$$d(\theta) = mr\left[\frac{I}{mr} + G(\theta).f(\theta)\right]$$

$$b(\theta) = -\frac{G(\theta)g(\theta)}{\frac{I}{mr} + G(\theta).f(\theta)}$$

Or the engine model may be written as:

$$\dot{\omega} = F(\omega, T_G - T_F, \theta)$$

(Eq 4.34)

Which define angular acceleration of shaft at zero load. The angular acceleration depends upon the operating point and system parameters.

4.4 Data Based Model

For a complex and uncertain system, it is difficult to establish a correct quantitative model. Uppal F. J, *et al* (2002, pp: 501-506) mentioned that when quantitative models of system are not readily available, a correctly trained neural network (NN) can be used as a non-linear dynamic model of the system. Wong P. K. *et al* (2009, pp: 55-72) characterized SI engine as a complex multivariable nonlinear function that is very difficult to determine. Wong also mentioned that a neural network is a universal estimator.

In production vehicles, the model of SI engine is assumed in the form of a number of lookup tables. The manipulating variables like intake air and spark position are measured and amount of fuel for the next ignition is decided by the entry of lookup table. Wong P. K. (2009, pp: 55-72) indicated that ECU of vehicles contain many lookup tables/ maps (like fuel map or ignition map). Uppal F. J, *et al* (2002, pp: 501-506) also indicated that grid based lookup tables are the most common nonlinear static model.

Optimal ECU setup maps are established by using engine models (e.g. MVM, DEM etc) and implemented in ECU through appropriate computer aided optimization methods. The car is finally required to go through a dynamo test for verification [Wong P. K. *et al* (2009, pp: 55-72)]. The vehicle maps are optimized for the new vehicles only. As the vehicle gets older, the model parameters start deviating from the reference parameters and the originally designed maps become sub-optimal. To maintain the vehicle performance over the entire life of vehicle, black box models with neural methods are being proposed.

Isermann R *et al* (2001, pp: 566-582) proposed local linear Neural Networks as a replacement of engine maps. In this method, he represented each neuron as local linear model with its own validity function. The whole neural network was therefore represented as a tree called "Locally Linear Model Tree" (LOLIMOT). For the system with n inputs $x_1, x_2, \ldots, x_n$, a Gaussian validity function $\Phi_i(x)$ was selected that determined the region of input space where each neuron is active. The output of model was formed by adding the contribution of all locally linear models as:

$$\hat{y} = \sum_{i=1}^{M} (w_{i0} + w_{i1}x_1 + \cdots + w_{in}x_n).\,\Phi_i(x) \qquad \text{(Eq 4.35)}$$

where w_{ij} are the parameters of i^{th} model.

LOLIMOT was trained in nested loop structure. Network structure was optimized in the outer loop that defined the number of neurons and partitioned the input space. The inner loop estimated the structure and parameters of local linear models. An effort was made that model remain valid over a wide operating range of one input variables and in small operating range of other variables.

The parameters of local linear model were estimated by a local weighted least square technique. The prediction errors of each model were weighted with the corresponding validity function. Each local model was estimated separately and the overlap between neighbored models was neglected. For multivariate, nonlinear dynamic processes it was assumed that the n inputs of neural network are function of p inputs of system as:

$$\underline{x}(t) = \left[u_1(t-1), \dots u_1(t-m), \dots., u_p(t-1), \dots, u_p(t-m), \hat{y}(t-1) \dots \hat{y}(t-m)\right]^T$$

(Eq 4.36)

i.e. each input of neural network is a function of inputs and its delayed versions. A block diagram of proposed general structure of LOLIMOT is shown in Figure 4.2.

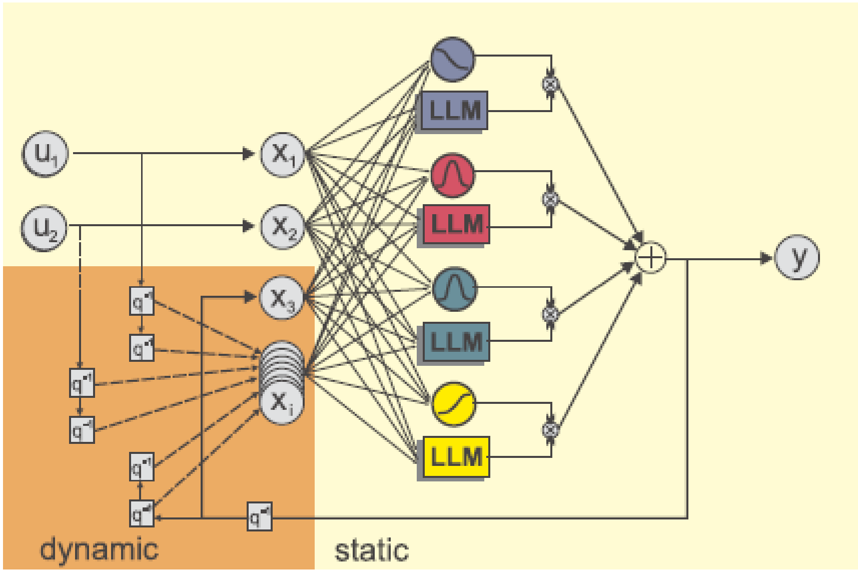


Figure 4.2 LOLIMOT net with external dynamics

Where LLM stands for local linear model and simple delay elements are expressed as q^{-1}.

Wong P. K. *et al* (2009, pp: 55-72) proposed an engine idle speed system modeling and control optimization using artificial intelligence. The proposed method is based on training the model using Least Square Support Vector Machine (LS-SVM). The training data was provided by Wong through the actual engine operation.

4.5 Hybrid Model

Hybrid system are characterized by the presence of both continuous time states and discrete time states in a model. A switched linear system is an important class of hybrid dynamical system, where dynamic system consists of a finite number of continuous time subsystems and logical rule that orchestrates switching between them [savskin]. The continuous states of system define the *state variables* of all of the continuous-time sub-systems and the discrete variable is the subsystem index.

A few references of modeling of SI engine as hybrid models are found in literature. Albertoni L *et al* (2003, pp:140-145) proposed hybrid model with four interacting subsystems: throttle valve, intake manifold, cylinders and crankshaft. Deligiannis V. F. *et al* (2006, pp. 2991-2996) represented SI engine as a hyper class of hybrid automata. The hyperclass automata are in general represented as a 12 tuple. Deligiannis V represnted SI engine by four states where the states were representing the suction, compression, power and exhaust strokes of engine. On the basis of simplifying assumptions a simple mathematical model was derived for the engine strokes. A summary of four states is provided below:

Induction or Suction Stroke

Instead of a constant pressure air suction stroke, a more accurate model was taken on the basis of literature review. The model of suction stroke taken by author was:

$$p = f(V^2)$$

Compression Stroke

The mathematical model for compression stroke was chosen as:

$$pV^{1.3} = constant$$

Power Stroke

It was assumed that combustion is instantaneous and the pressure rises abruptly. The governing model of power stroke was assumed as:

$$pV^{1.48} = Constant$$

Exhaust Stroke

An abrupt pressure drop was assumed in exhaust stroke. Instead of assuming exhaust at constant pressure, the goverrning model of exhaust stroke was assumed as:

$$p = f(-V^2)$$

The switching logic was defined on the basis of volume taking extreme values i.e. when $V = V_{max}$ the state transition occur from suction to compression or from power stroke to exhaust stroke. Similarly when $V = V_{min}$ the state transition occur from exhaust to suction stroke or from compression to power stroke. The transition between states was assumed deterministic and cyclic as shown in Figure 5.3. The volume of air inside the cylinder defined guard condition for the switching of states.

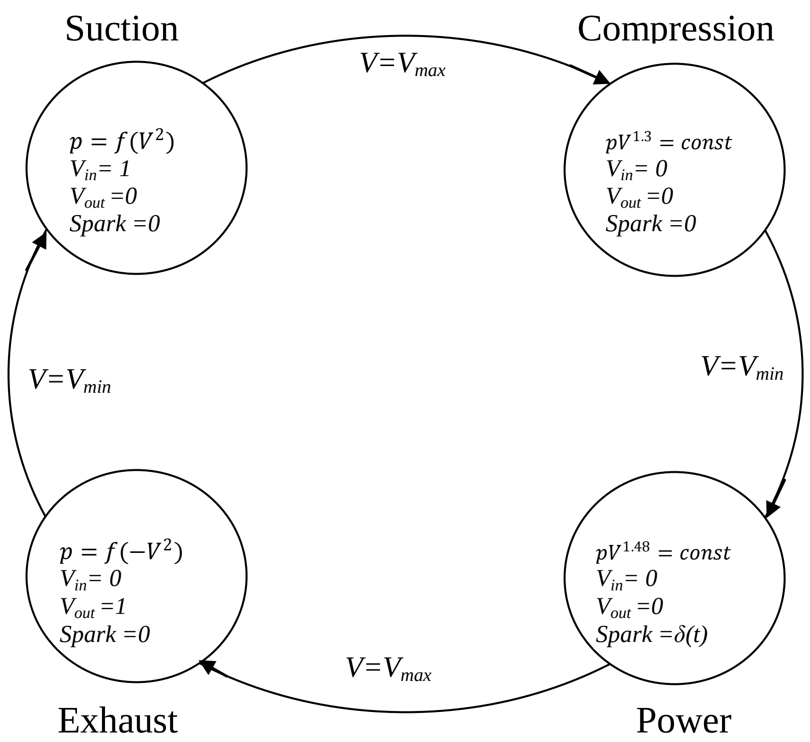


Figure 4.3: Engines Automaton with four States .

The hybrid models of SI engine proposed in literature are either based on the engine system consisting of subsystems like throttle, cylinder, manifold and crankshaft or are based on the thermal processes occurring in cylinders i.e. suction, compression, power and exhaust. In both these approaches the continuous dynamics of subsystems is quite different from each other. The hybrid model proposed in this research [Rizvi M. A *et al* (2009, pp: 1-6), Rizvi M. A. *et al* (2010)] has adopted quite a different strategy by considering the four cylinders of engine as four different subsystems of hybrid model. The continuous dynamics of subsystem is governed by the movement of piston inside the cylinder and switching of subsystems is defined by the position of crankshaft. The basic advantage of this strategy is that under ideal conditions, all the four subsystems of hybrid model would be identical which simplify the calculations.

When the engine is operating normally, all the models can be used to study the engine behavior in their respective domain. If however the switching sequence of engine gets disturbed due to some faults or by mistakenly interchange cylinder ignitor, then MVM cannot explain the resulting engine behavior. Hybrid model can however be applied to study this behavior also.

Summary

This chapter provided a comprehensive survey of different mathematical models observed in literature to represent the spark ignition engine. The basic philosophies behind the development of those models were discussed. The strong and weak features of models were elaborated. It was concluded that MVM is predominantly used for the application of controller design due to its accuracy, better physical insight and simplicity when compared to DEM model of SI engine. The major weakness of MVM is that it operates on the basis of average values of engine variables in a complete ignition cycle. It is concluded that the model is not a suitable option for detecting the misfire fault, as the fault detection need analysis within an ignition cycle. It was also discussed that DEM model is more accurate and provide insight of engine operation even within an ignition cycle. The computational load of DEM is however sufficiently high. The literature survey also provided a few SI engine models based on kinematic behavior of SI engine that analyze the forces acting on engine piston, when fuel is burnt inside the chamber. The mentioned kinematic models are based on the physical relations between different engine variables and are derived using basic laws of physics.

Some Data Based Models of SI engine found in literture were also discussed in this chapter. LOLIMOT is using artificial neural network to train model parameters to approximate the observed engine response. Similarly in another approach, the training of model is carried using SVM.

Finally it was mentioned that hybrid approach to represent the approximate response of SI engine is introduced recently in literature where the independent spark events are taken as the discrete events and the crankshaft speed is considered as the continuous variable. The integration of discrete and continuous variables to represent the engine response is used as the basis of hybrid modeling.

Chapter 5

HYBRID MODEL FOR SI ENGINE

Although model based methods are applicable under all operating conditions, however for complex systems like SI engine the chances of development of an accurate yet simple model are very little. The available models of SI engine are only too complex to become useful for misfire fault detection application. Due to this non-availability of appropriate model, misfire fault detection in SI engine is predominantly carried out using data based technique. Most of the data based methods used some heuristic guide line and worked directly on experimental data which is obtained by engine operating under specific operating conditions. Although the heuristic guideline successfully achieved the requirement of fault detection, however the lack of concrete theoretical model behind the method resulted in some severe short comings. Sood A. K *et al* (1985, pp: 301-307) presented both model based technique and pattern recognition methods for detecting misfire fault and claimed that it is advantageous to use physical reasoning for the development of classification rules. Angeli C (2004, pp: 12-30) indicated that heuristic based expert system can guide efficient diagnostic procedure but they lack generality. He clearly indicated that rule based approaches poorly handles the novel situations and mentioned that the main weakness of rule based methods can be eliminated using model based methods that can even handle the unexpected cases not covered by the heuristic based methods. In addition to the major weaknesses of these methods mentioned above, some other shortcomings of these methods include:

- To validate the fault detection method by simulation, reference data cannot be generated even assuming ideal conditions.
- Inability of method to analytically explore the performance of fault diagnostic algorithm against e.g. time delay between initiation and detection of fault and percent isolation to one line replacement unit, defined by George V. [2006, pp: 362-363].

- Inability of method to explore the properties of process variables. As an example an appropriate model enables us to identify that the variable of interest is Markov/ Gaussian. The information about properties of process variable provides us sufficient insight to identify some new and stronger prediction algorithms to detect fault.

The data based methods developed under heuristic guideline were based either on the basis of black box models like artificial neural network or on the basis of signal shape defined by domain experts on the basis of their observations. These methods suffer from the following shortcomings:

- Methods based on artificial neural network lack physical insight of process.
- It is not possible to study all possible conditions to identify different signal shapes for training the network. Also characteristics of systems change over their life time, hence the trained neural network, or data based features need to be modified.
- Some of the methods are applicable at the operating point only. The method can however be applied at other operating point after necessary changes e.g. for fault isolation the method of correlation analysis described in Section 3.3.2 needs correlation with the waveforms (some of which are shown in Fig. 3.3) that correspond to experimental observation under different misfire conditions. When the method would be applied at some other operating point, the new experimental observations would be needed because the frequency of oscillation would change with operating point. Also different results may also be observed with the aging of engine or even when experiment is conducted on another SI engine of different model or make.

Some of the mentioned problems can however be avoided if a heuristic affect is identified in the system response, that can be used for fault detection purpose under all operating conditions. A simplified model could be developed to illustrate that effect. For SI engine crankshaft speed fluctuations are a main affect that exist in steady state response under all operating conditions but the frequency and amplitude of fluctuations vary with the operating point and a simple comparison of waveform would not work. This effect is however not explained by MVM. Although DEM can explain the speed fluctuations but the computational complexity of DEM is very high and study of statistical variations of peaks of speed fluctuations is very difficult. It is therefore desired to develop a new simplistic model that can not only be used to study the crankshaft speed fluctuations but could also analyze the statistical variation of peaks of speed fluctuation signal.

The model properties could be studied to identify the affects that are present under all operating conditions and could then be explored to identify the properties that can be used for fault identification under all operating conditions. This approach can be considered as an integrated data and model based approach, where model would be used to provide a heuristic guideline for the development of data based algorithm. The approach of integration of data and model based methods for the fault diagnostic applications is a fairly recent approach.

Roemer M *et al* (2006, pp: 707-715) mentioned that an integrated physical and stochastic modeling approach suits well for prognosis applications in the presence of uncertainties like load variations, parameter variations etc. He indicated that data based model like neural network or probabilistic methods can use the results from the physics based model. He also mentioned that a combined model based and feature based approach provide full prognostic ability over the entire life of the component under observation and provide information to properly plan the maintenance of component during system overhauling.

The basic modeling objective can therefore be identified as an aid that provides a heuristic guideline for data based approach. In this strategy, the developed model is simple so that the statistical properties of model input variations could also be studied. For the application in hand, the accuracy of model can therefore be compromised as long as the model provides correct heuristic guide line to the fault diagnostic method. The statistical properties of model variable would be carried out to identify that peaks of crankshaft speed fluctuation signal are Gaussian and Markov. The establishment of Markov property would provide us a heuristic guideline to transform the model outputs to some discrete states and use Markov chains to predict the future behavior of those states to identify the faults. The basic modeling objectives in this research study can therefore be summarized as:

- The model should be simpler so that statistical properties of some model variable could be explored to identify that the variable is Markov

- Identify some property of model output that remains valid under all engine operating conditions.
- Development of finite number of discrete states from model that can be modeled as discrete Markov chains.
- Validation of model using experimental results.
- Generate data for simulation under different operating conditions for testing of algorithm and study the features of algorithm that could not be verified experimentally due to absence of appropriate experimental facility.

Although the model proposed in this thesis can be used to develop misfire detection methods using a number of tools of FDI and DX community, *the study in this thesis is restricted so that the heuristic guideline provided by the model can be used to apply Markov Chains for misfire fault detection.*

The application of Markov Model in the area of fault detection is elaborated in Section 3.8. The literature survey clearly indicates that fault diagnosis algorithms based on Markov models were used for fault prediction by Morgan I. *et al* [2009, pp: 1774-1781], for path prediction by Yu *et al* [2006, pp: 374-379] and for traffic forecasting by Sun S. *et al* [2004, pp: 437-441]. The prediction/ forecasting of fault in a system correspond to early fault detection in system. The works available in literature clearly indicate that Markov chain is a suitable technique for early fault detection. The application of Markov model in the area of fault diagnosis in SI engine is not explored widely. The basic reason of this lack of interest is the unavailability of appropriate mathematical models in which the variables of interest could be represented as a set of discrete states satisfying the Markov condition.

Smith F. S *et al* [2004, pp: 649-663] used a General Diagnostic Engine (GDE) for fault detection using belief revision method. In the proposed method, a large system was represented by a set of interacting components each with a fixed set of possible behaviors that are independent from each other. A variable of interest could however be generated by a number of different routes that encounter different model fragments. The condition of independence of behaviors was termed mandatory by the author. The possible set of behaviors consist of response of component under different operating conditions e.g. under no-fault condition, faulty condition etc. He referred the behavioral descriptions as model fragments. Each model fragment had a *belief* attached to it that indicates how close a particular fragment is to the actual behavior of component. The diagnostic method model was simulated and the variable of interest was analyzed by processing the inputs through various model fragments. The belief of a model was updated using "*confirm set*" or a "*conflict set*". A confirm set was generated for model when two identical values of variable of interest could be generated by simulating in two different ways. In case of non-identical values a conflict set was generated. The "confirm/ conflict sets" were then used to generate the penalizing and rewarding evidences. Markov chain was then used to process these evidences after necessary conversion to a Markov matrix.

Although the scope of proposed method of Smith is different from the method being proposed in this thesis, however the conditions imposed on the system by Smith before application of Markov chains should be given a due consideration. Some of the conditions are:

- Division of whole system into a set of interacting components exhibiting specific behavior
- The behavior of different components be independent from each other
- The processing of variable of interest to generate the states like "*confirm / conflict"* or *"rewarding / penalizing"* evidences.

The study of different modeling approaches of SI engine indicates that MVM is based on the physical modeling of engine processes on the basis of modeling of fluid flow and torque produced at the output. The geometry/ mechanical design of SI engine is however not considered in modeling so the division of system into a number of interacting components using MVM is not possible.

DEM defined the system in states. DEM considered not only each cylinder as independent physical entity in model but also considered the thermodynamic cycle of each cylinder independently. In this way suction, compression, expansion and exhaust strokes are considered separately for each cylinder. The behavior of each stroke is studied on the basis of principles of physics. The resulting model is therefore highly complex but is capable to detect the misfire fault as described by Guzzella L. (2004, pp: 23)]. The complexity of DEM model however restricted its application on wide scale and the model was used only by limited research community, mostly for control applications only.

The proposed hybrid model can be considered as a simplification of DEM in which only the power stroke of system is modeled. The simplified model is then validated by simulation and experiments. The proposed model was analyzed for the conditions proposed by Smith as well as the validity of Markov model before application of algorithm based on Markov chain.

The condition of dividing of system into interacting components is quite natural in SI engine where the engine system can be considered as a number of cylinders interacting with each other. However all the cylinders are physically coupled with each other through a common crankshaft mechanism, the condition of independence of interacting components is difficult to establish and needs to be studied. Another problem in the application of Markov chain is to establish the Markov condition for the variables of interest.

The difficulties in using the available SI engine models for use in a diagnostic algorithm using Markov chain instigated for the development of a new hybrid model and study its properties for the application of diagnostic method. As the proposed model is hybrid in nature, it is appropriate to study the properties of general hybrid system. Section 5.1 provides an orientation of hybrid system and its mathematical representation. The hybrid modeling of SI engine is presented in section 5.2 and the statistical properties of model are provided in section 5.3 to establish the condition that response of components are independent and define the variable of interest in model which is Markov.

5.1 Hybrid Systems

Lunze J. mentioned that hybrid dynamical system has emerged out to become a major research topic in last decade of twentieth century [Engell S *et al* (2002, pp: 3)]. Many engineering system are represented by hybrid model. Although some hybrid models for SI engine are found in literature, the field is not well established. Before starting hybrid modeling of SI engine it would however be beneficial to define the hybrid system and study its properties.

Liberzon D. (2003, pp: 3) defined the hybrid systems as dynamic systems whose output is defined by an interaction between continuous and discrete dynamics. Many man-made systems could be represented well as a hybrid system. The hybrid nature of model arises either due to the switching of controller in different ***modes*** of process (e.g. gain scheduling applications) or the dynamics of process is switched while the controller remain unchanged. Schutter D. *et al* , (2003, pp. 8-9) described many possible modeling techniques for a hybrid system; some of which are:

- Timed or hybrid Petri-nets
- Hybrid automata
- Mixed logical dynamic models
- Real-time temporal logics
- Timed communicating sequential processes
- Switched bond graphs

Diverse nature of modeling techniques in development of hybrid system is due to different communities applying it in their respective domains. Liberzon mentioned that different communities have contributed to the development of hybrid systems according to their requirements e.g. researchers in computer science concentrate on studying the discrete behavior of system assuming a simpler continuous dynamics. The community of control systems on the other hand stressed on the properties of continuous system with simpler state switching. Control community referred these systems as switched systems. The switching sequence in these systems is defined either as a function of states or as a function of time [Liberzon D. (2003, pp: 3), Zhendong S. (2005, pp:16-18)].

Hybrid modeling approach was used for other automotive applications by many researchers like Torrisi F. D *et al* [2004, pp: 235-249] and Giorgetti N. *et al* [2006, pp: 499-506], Balluchi A [2000, pp: 888-912] etc. The application of hybrid modeling for fault detection of SI engine is however not extensively investigated.

5.1.1 Switched Linear System

The class of hybrid system where all the subsystems are linear time-invariant systems is commonly termed as switched linear systems [Zhendong S. (2005, pp:3)]. Tabuada P. (2009, pp: 3-4) described a notation of a set tuple $< X, X_0, u, \rightarrow, Y, H >$ to represent the hybrid systems; where:

- X represent the state variables of system,
- X_0 is the set of initial states,

- u represent the system inputs,
- $\rightarrow$ represent transition relations,
- Y represent system output
- H is the output map

Hybrid systems exhibit a number of properties.

5.1.2 Properties of Switched Linear System

- A switching system is said to exhibit Zeno Behavior if there exist two states x_0 and x_1 such that starting from state x_0 the system need infinite number of jumps to reach the state x_1 in a finite time t. Zeno behavior is not necessarily exhibited by all switched systems.
- A switching path is said to be well defined on $[t_1, t_2]$, if it is defined in $[t_1, t_2]$ and for all $t \in [t_1, t_2]$, both $lim_{s\uparrow t}\theta(s)$ and $lim_{s\downarrow t}\theta(s)$ exist, and it has only finite jump instants in any finite time interval $[t_1, t_2]$. The possibility of well definedness excludes the property of Zeno behavior.
- Stability of switching system is defined not only by the system dynamics but also depend on the switching sequence. Liberzon D. (2003, pp:19) indicates that given two stable systems and discrete dynamics to switch the two systems to get an output. A switching dynamics can be defined that would result in an unstable system. Also a switching sequence can be defined that could stablize a switched system with a stable and an unstable sub-systems. The stability of switching systems can be determined by finding a common Lyapunov function for all the sub-systems. The stability of Switched and hybrid systems were studied by Branicky M. S. (1998, pp: 475-482) and are discussed in Schutter D. *et al* , (2003, pp. 67-82)
- A hybrid/ switching system exhibit some oscillations (similar to chattering) about its operating point.

An SI engine exhibit both continuous and discrete dynamics so it can be represented as a hybrid system. The hybrid model proposed in this chapter is developed with the basic objective of acting as an aid to misfire fault detection. Since fault detection can be carried out in the controlled environment, a simplified steady state model with some fixed parameters like fixed throttle position and constant load conditions etc can be considered. Such a steady state model could be used for the development of relatively simple fault diagnostic algorithms.

5.2 Hybrid Model of SI Engine

Although the representation of SI engine as a hybrid model is already present in literature, the main difference of the approach presented in this thesis is the manner in which the continuous states of model are being represented. The hybrid model presented by Deligiannis V. F *et al* (2006, pp. 2991-2996) assumed the model of four engine processes i.e. suction, compression, power and exhaust as four continuous sub-systems. Similar continuous systems are also considered in DEM that can also be considered as a hybrid model. In this model, each cylinders of engine is considered as independent subsystem that takes power generated due to the burning of air fuel mixture as input and movement of piston in engine cylinder is considered as the output. These sub-systems are represented as linear systems and complete SI engine is considered as a collection of subsystems. These subsystems are working coherently to produce the net engine output. The proposed hybrid model of SI engine can be regarded as a switched linear system. Although an SI engine is a highly nonlinear system, for certain control applications a simplified linear model is used. Lee M. *et al* (2006, pp: 637-644) mentioned that modeling assumption of constant polar inertia for crankshaft, connecting rod and piston assemblies to develop a linear model is a reasonable assumption for a balanced engine having many cylinders. The modeling of sub-systems of proposed hybrid model would be performed under steady state conditions, when the velocity of system is fairly constant. Also the time in which the sub-system give its output is sufficiently small. A linear approximation for modeling of sub-system can therefore be justified. Similar assumption of locally linear model is made by Isermann R *et al* (2001, pp: 566-582) in LOLIMOT structure. The continuous cylinder dynamics is therefore represented by a second order transfer function with crankshaft speed as output and power acting on pistons of cylinder due to fuel ignition as input.

A continuous dynamic model of these sub-systems would be derived in this chapter. The timing of signals to fuel injectors, igniters, spark advance and other engine components is controlled by Electronic Control Unit (ECU) to ensure the generation of power in each cylinder in a deterministic and appropriate order. The formulation of hybrid modeling of sub-systems would be carried under the following set of assumptions:

Modeling Assumptions

1. Engine is operating under steady state condition at constant load.
2. Air fuel ratio is stoichiometric.
3. Air fuel mixture is burnt inside engine cylinder at the beginning of power stroke and energy is added instantaneously in cylinder resulting in increase in internal energy. This internal energy is changed to work at a constant rate and deliver energy to a storage element (flywheel).
4. At any time instant only one cylinder would receive input to become active and exerts force on piston and other cylinders being passive due to suction, compression and exhaust processes contribute to engine load torque.
5. All the four cylinders are identical and are mathematically represented by the same model

The switching logic can be represented as a function of state variables of systems.

5.2.1 Framework of Hybrid Model

The framework of Hybrid model for a maximally balanced SI engine with four cylinders is represented as a 5-tuple model $< \mu, X, \Gamma, \Sigma, \phi >$. The basic definition of model parameters is given below.

- $\mu = \{\mu_1, \mu_2, \mu_3, \mu_4\}$ where each element of set represents active subsystem of hybrid model.
- $X \in R^2$ represents the state variable of continuous subsystems, that would be defined when model is developed for subsystems, where the vector X consists of velocity and acceleration.
- $\Gamma = \{ M \}$ is a set that contains only a single element for a maximally balanced engine. M represents state space model of all subsystems and is assumed to be linear, minimum phase and stable. The model equation is derived in the next section. The model can be defined in state space as:

$$\dot{X}(t) = AX + BU \qquad \text{(Eq 5.1)}$$

$y(t) = CX + DU$

Where

$U \in R,\ A \in R^{2\times 2}\ ,\ B \in R^{2\times 1},\ C \in R^{1\times 2},\ D \in R$

- $\Sigma: \mu \rightarrow \mu$ represents the generator function that defines the next transition model. For an IC engine, the piston position has a one to one correspondence with crankshaft position during an ignition cycle. The generator function is therefore defined in terms of crankshaft position as:

$$\Sigma = \begin{cases} \mu_1 & 4n\pi \leq \int \dot{\theta}_1 dt < (4n+1)\pi \\ \mu_2 & (4n+1)\pi \leq \int \dot{\theta}_1 dt < (4n+2)\pi \\ \mu_3 & (4n+2)\pi \leq \int \dot{\theta}_1 dt < (4n+3)\pi \\ \mu_4 & (4n+3)\pi \leq \int \dot{\theta}_1 dt < (4n+4)\pi \end{cases} \qquad \text{(Eq 5.2)}$$

---

where n=0,1,2,… and $\int \dot{\theta}_1 dt$ represents instantaneous shaft position that identifies the output of generator function.

- $\phi: \Gamma \times \mu \times X \times u \rightarrow X$ defines initial condition for the next subsystem after the occurrence of a switching event, where u represents input to subsystem. Figure 5.1 shows the subsystems and switching sequence of proposed SI engine hybrid model.

5.2.2 Modeling of Sub-system

A subsystem/cylinder is *active* when it contributes power to system i.e. during power stroke. When a sub-system is active its output is defined by the dynamic equations of system and its output during its inactive period is defined by its storage properties. The output of a sub-system provides initial condition to the next sub-system at the time of switching. All the subsystems are actuated sequentially during an ignition cycle. The cyclic actuation of subsystems is represented as a graph in Figure 5.1. The total output delivered by the system during complete ignition cycle would be the vector sum of outputs of all subsystems during that ignition cycle.

If T is the period of ignition cycle and *u(t)* is the input to system at time *t* within an ignition cycle and $u_i(t)$ is the input of i^{th} subsystems; by assumption 4:

$$u_i(t) = u(t) \qquad when \quad \frac{(i-1)T}{4} < t < \frac{iT}{4} \quad , i = 1,2,3,4$$
$$u_i(t) = 0 \qquad otherwise \tag{Eq 5.3}$$

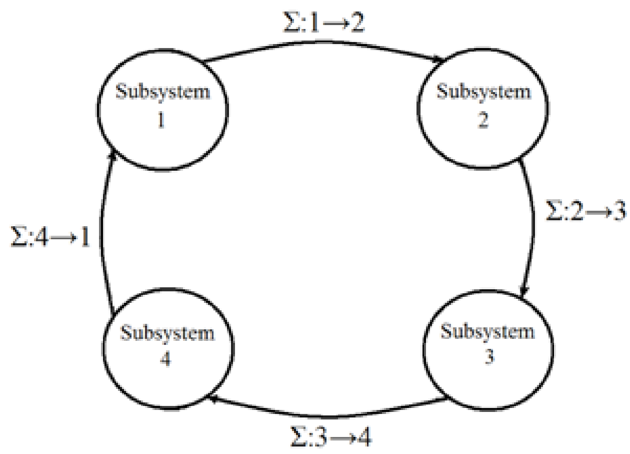


Figure 5.1 : Switching of subsystems *(Adopted from Rizvi (2009, pp. 1-6))*

5.2.2.1 Modeling of Sub-system

Franco *et al* (2008, pp: 338-361) used mass-elastic engine crank assembly model for real time brake torque estimation. In this representation of SI engine each cylinder is represented by a second order mass spring damper as shown in Figure 5.2.

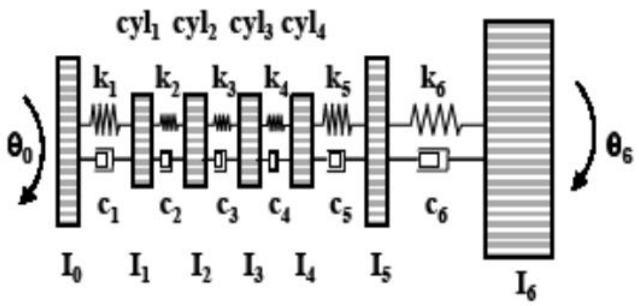


Figure 5.2: Spark ignition engine representation *(Adopted from Franco et al (2008, pp. 338-361))*

Consider δQ amount of energy added in system by burning air fuel mixture. The instantaneous burning of fuel increase the internal energy δU in cylinder chamber.

$$\delta U = \delta Q \tag{Eq 5.4}$$

At ignition time, energy is added instantaneously in engine. This will increase internal energy of system. A part of this internal energy is used to do work and rest of the energy is drained in coolant and exhaust system. If internal energy change to work with constant efficiency η_t then work δW is given by the energy balance equation as:

$$\delta W = -\eta_t \, \delta U \tag{Eq 5.5}$$

$Using\ equation\ (5.4)\ we\ get$

$$\delta W = -\eta_t \, \delta Q$$

If p is pressure due to burnt gases then work done during expansion stroke is given by:

$$W = \int_{V1}^{V2} p dV \tag{Eq 5.6}$$

where V_1 and V_2 are initial and final volume of cylinder during expansion. For adiabatic expansion:

$$pV^{\gamma} = k_1 \tag{Eq 5.7}$$

where k_1 and γ are constant. Hence Eq 5.6 becomes

$$W = \int_{V1}^{V2} k_1 V^{-\gamma} dV$$

$$W = k_1 \frac{V_2^{-\gamma+1} - V_1^{-\gamma+1}}{-\gamma+1} \tag{Eq 5.8}$$

Consider that the closed end of the piston to be *origin* and x is a continuous variable representing the instantaneous piston position with respect to the origin. The piston always moves between two extreme positions x_t and x_b where x_t represent piston position at Top Dead Center (TDC) and x_b represent piston position at Bottom Dead Center (BDC). If the surface area of piston is A, and it moves a small distance δx from its initial position x, where δx is constant and can be chosen arbitrarily small, then using Eq 5.8 work done can be expressed as:

$$\delta W = k_1 \frac{[A(x+\delta x)]^{-\gamma+1} - [Ax]^{-\gamma+1}}{-\gamma+1} \qquad \text{(Eq 5.9)}$$

$$\delta W = k_1 \frac{A^{-\gamma+1}}{-\gamma+1}\left[(x+\delta x)^{-\gamma+1} - x^{-\gamma+1}\right]$$

$$\delta W = \frac{k_1 A^{-\gamma+1}}{-\gamma+1}\left[x^{-\gamma+1}\left(1+\frac{\delta x}{x}\right)^{-\gamma+1} - x^{-\gamma+1}\right]$$

$$\delta W = \frac{k_1 A^{-\gamma+1} x^{-\gamma+1}}{-\gamma+1}\left[\left(1+\frac{\delta x}{x}\right)^{-\gamma+1} - 1\right]$$

Expanding using binomial series and neglecting higher powers of δx and simplifying:

$$\delta W = k_1 A^{-\gamma+1} x^{-\gamma}\,\delta x$$

Therefore from Eq 5.5

(Eq 5.10)

$$\delta Q = -\frac{k_1 A^{-\gamma+1} x^{-\gamma}\,\delta x}{\eta_t}$$

In deriving the model for sub-systems, each cylinder of SI engine is treated as a second order system as used by Franco *et al* (2008, pp: 338-361) and shown in Figure 5.2. Consider if F is the applied force by the burnt gases, m is the mass of engine moving assembly (piston, connecting rod, crankshaft and flywheel), coefficient of friction is k_2 and coefficient of elasticity is k_3, then net force acting on piston is given by:

$$m\frac{d^2x}{dt^2} = F - k_2\frac{dx}{dt} - k_3x$$

$$m\frac{d^2x}{dt^2} + k_2\frac{dx}{dt} + k_3x = F \qquad \text{(Eq 5.11)}$$

Net work done by the expanding gases against the load, friction and elastic restoring forces when piston moves by a small distance δx would be given as:

$$\left[m\frac{d^2x}{dt^2} + k_2\frac{dx}{dt} + k_3x\right]\delta x = \delta W \qquad \text{(Eq 5.12)}$$

Using Eq 5.10 above equation becomes

$$\left[m\frac{d^2x}{dt^2} + k_2\frac{dx}{dt} + k_3x\right]\delta x = k_1A^{-\gamma+1}x^{-\gamma}\delta x$$

The displacement δx can be chosen constant and arbitrarily small. As the piston moves, the volume inside the combustion chamber increases resulting in the reduction of instantaneous pressure on piston. Instantaneous power is therefore a function of piston position. Instantaneous power delivered by the engine would be calculated by differentiation as:

$$\left[m\frac{d^3x}{dt^3} + k_2\frac{d^2x}{dt^2} + k_3\frac{dx}{dt}\right]\delta x = -k_1\gamma A^{-\gamma+1}x^{-\gamma-1}\frac{dx}{dt}\delta x$$

$$m\frac{d^3x}{dt^3} + k_2\frac{d^2x}{dt^2} + k_3\frac{dx}{dt} = -k_1\gamma A^{-\gamma+1}x^{-\gamma-1}\frac{dx}{dt} \qquad \text{(Eq 5.13)}$$

Writing differential Eq 5.13 in terms of velocity *v* as:

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = -\gamma\eta_t\frac{k_1A^{-\gamma+1}x^{-\gamma}\delta x}{\eta_t}\frac{v}{x\delta x}$$

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = \gamma\eta_t\delta Q\frac{v}{x\delta x} \qquad \text{(Eq 5.14)}$$

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = \frac{\gamma\eta_t v}{x}\frac{\delta Q}{\delta t}\frac{\delta t}{\delta x}$$

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = \frac{\gamma\eta_t v}{x}P(x)\frac{1}{v}$$

Assuming that crankshaft speed is proportional to the speed of piston inside the cylinder, Eq 5.14 represents a model of crankshaft speed when energy is added in one of the cylinder of SI engine by the ignition of fuel. The model is however nonlinear on account of presence of x in the denominator on the right side of differential equation.

5.2.2.2 Model Linearization

SI engine is a highly nonlinear system. In hybrid modeling, the time of activation of subsystems is very small. Also under under steady state conditions, the velocity of engine is fairly constant hence a linear approximation of engine subsystems can be justified. The model derived in earlier section is now linearized to form a switched linear model. The validity of linear model is only at the operating point. As x can never be zero, so the function is smooth and can be linearized at TDC. If the igniting fuel adds the power $P(x)$ to a cylinder when piston is at position x, the dynamics of system at TDC would be described as:

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = \frac{\gamma\eta_t}{x}P(x) \qquad \text{(Eq 5.15)}$$

Linearizing the system at TDC ($x = x_t$) under steady state condition, and assuming that whole power is added in the cylinder instantaneously when the cylinder is at TDC, Eq 5.15 becomes:

$$m\frac{d^2v}{dt^2} + k_2\frac{dv}{dt} + k_3v = \frac{\gamma\eta_t}{x_t}P(x_t) \qquad \text{(Eq 5.16)}$$

In simulations, *P(x)* can be taken as a narrow pulse or a triangular wave, assuming that when system receives input, it deliver power at constant high rate for a short interval of time and thereafter the delivered power would be negligible. Since shaft speed is also constant at the start of each ignition cycle, therefore right hand side of Eq 5.16 becomes constant and expression becomes a linear differential equation.

5.2.3 Model Properties

This section correlate some general observations of SI engine with the properties of hybrid model and some results would be established for further analysis.

- Even under steady state conditions, SI engine exhibit a slight oscillatory component in crankshaft angular speed. These oscillations can be attributed due to switching of hybrid system.
- The number of ignitions in engine cylinders in a finite time is also finite. This indicates that SI engine do not exhibit Zeno behavior.
- To increase the engine speed, the ignition event would have to occur more frequently. The system can be steered from any given state at any arbitrary time to some other state by changing the switching path in a finite interval of time. SI engine can therefore be considered as a well posed system.
- When all the four cylinders are identical, the transfer function of all cylinders is identical and a common Lyapunov function can be found to ensure the stability of engine system.

The mentioned observations correlate well with the properties of hybrid model providing some justification of using a hybrid model for SI engine. This section would study the properties of engine that can be deducted from hybrid model. The properties would be studied in the form of some propositions and Lemmas applicable to hybrid model of SI engine. Keeping in mind the basic working and operation and working of SI engine, the mentioned properties seems to be trivial; however most of the mentioned properties could not be explained on the basic of MVM. These properties would be used in next sections for the development of misfire detection algorithms for SI engine.

When same air input is provided to all cylinders and engine is operating under steady state conditions, the model exhibits the following properties.

5.2.3.1 Proposition 1

When engine is running with constant speed, input to engine system is a periodic impulse train with period T/4 where T is the period of ignition cycle. This is because input impulse is given after angular displacement of π and under constant speed assumption. The same time would be needed for each displacement.

5.2.3.2 Lemma 1

Under no misfire condition the system output exhibits a periodic AC component with period T/4 where T represents the period of ignition cycle (two revolutions).

Proof:

When all subsystems are identical and represented by a linear model, then a periodic input with period T/4 would produce a periodic output with same fundamental frequency.

5.2.3.3 Lemma 2

When a cylinder misfires, the output of system exhibits a periodic AC component with period T.

Proof:

Misfire can be considered as the loss of one of every fourth impulse of input signal. The input signal is therefore periodic with period T rather than T/4 and output also exhibits fundamental frequency T.

5.2.3.4 Lemma 3

For a four cylinder engine, with no misfire fault condition, the output contains four identical peaks.

Proof:

All subsystems are represented by stable, LTI minimum phase second order system that exhibits a single peak in output against each impulse input.

5.2.3.5 Lemma 4

In steady state conditions with fault in i^{th} event, no peak would be observed due to input of i^{th} subsystem.

Proof:

Absence of impulse at ith place in input signal would result in the loss of corresponding peak.

The results of lemma 1, 2, 3 and 4 can be observed from simulation results discussed later.

5.2.3.6 Defintion 1

The system is said to be in steady state when the net change in system output *v*(t) in one complete ignition cycle is zero.

$$v(t+T) = v(t) \qquad \text{(Eq 5.17)}$$

5.2.3.7 Theorem 1

Under steady state and no fault condition when same input is given to identical subsystems (maximally balanced cylinders), the response of each subsystems would be independent:

Proof: If *u* is the input to a subsystem, *v(0)* is initial condition and *h(t)* is the impulse response of a subsystem then the output of second subsystem i.e. at time *t* where $T/_4 < t < {}^{2T}/_4$ is given by:

$$v(t) = v(0) + \int_0^{\frac{T}{4}} h(t-\tau)u(\tau)d\tau + \int_{\frac{T}{4}}^{t} h(t-\tau)u(\tau)d\tau \qquad \text{(Eq 5.18)}$$

By Lemma 1 the output signal is periodic with period $\frac{T}{4}$ therefore

$$v\left({}^{T}/_{4}\right) = v(0)$$

$$\Rightarrow v(0) + \int_0^{\frac{T}{4}} h(t-\tau)u(\tau)d\tau = v(0) \qquad \text{(Eq 5.19)}$$

And Eq 5.18 becomes:

$$v(t) = v\left(T/_{4}\right) + \int_{\frac{T}{4}}^{t} h(t-\tau)u(\tau)d\tau$$

Hence during the activation time of second subsystem, system output depends only on the input *u(t)* and impulse response of second subsystem and is independent of response of first subsystem. Similarly, it can be proved that under steady state conditions, responses of all cylinders are independent.

5.2.3.8 Proposition 2

For an EFI engine, air intake in cylinders is measured and fuel proportional to amount of air intake is sprayed in it. Therefore power input to system under steady state conditions is proportional to the amount of air intake.

The mentioned model properties can be summarized in the form of four corollaries

5.2.3.9 Corollaries

- Four peaks would be observed in one ignition cycle of a four cylinder SI engine. (Lemma 3)
- Amplitude of four observed peaks represents four independent events. (Theorem I)
- Crankshaft speed is proportional to input power. (being input and output of linear model of subsystems)
- Crankshaft speed is proportional to amount of intake air. (By Corollary 3 and Proposition 2)

5.2.4 Model Input Estimation

The input to the model is the power generated inside the cylinder as a result of ignition. It is assumed that power operating on piston is coming from two sources i.e. by the ignition of fuel and by the power supplied by the engine rotating assembly due to inertia. In case of misfire, the power due to inertia of rotating assembly will maintain the movement of piston but the Power due to ignition of fuel is absent. Power can be defined as the product of force acting on piston of a cylinder and piston velocity. If F is the force acting on engine piston and v is the piston velocity, then power P acting on piston can be defined as:

$$P = Fv$$
$$P = p.A.v \qquad \text{(Eq 5.20)}$$

Where p is the pressure inside the cylinder, A is the surface area of piston which is known.

A typical curve of cylinder pressure signal as a function of crankshaft angular position given by Isermann R. (2001, pp: 566-582) is shown in Figure 5.3. Input to cylinder is assumed to be a pulse. To arrive on a rough estimate of pulse amplitude, cylinder pressure signal is approximated as a pulse starting at TDC and remain for 60˚ rotation. The peak of cylinder pressure signal is considered as 25 bars (assumed as a typical value). If engine is running at 15 revolutions per second i.e. idle speed, and cylinder stroke is 75mm, then average speed of piston can be easily estimated. This pulse would be provided once in each ignition cycle i.e. in 720˚. The time to traverse the complete stroke is 1/30 seconds or nearly 0.03 seconds. The average power provided by the fuel can now be estimated as:

$$Power = \left(\frac{2500000}{12} \times pi \times .075 \times .075/4\right) \times \left(\frac{.075}{.03}\right)$$
$$Power = 2301\ Watt = 3.1\ hp \qquad \text{(Eq 5.21)}$$

The net power contribution of four cylinders would be nearly 12 hp which is quite a reasonable power contribution under idle conditions. The plausibility of result provided basis to test the approach in simulation. The simulation results are provided in chapter 7.

By (Eq 5.20) the steady state input power is defined by pressure inside the cylinder, piston area and piston speed. The only unknown variable is the cylinder pressure that can be estimated using observer or an estimator. One such technique of cylinder pressure estimation proposed by Yaojung S. and Moskwa J. (1995, pp: 70-78) is described briefly in section 3.2.2.1. The proposed method or some similar method that takes only crankshaft speed signal as input for pressure estimation can be used.

Guzzella L. (2004, pp: 165) has provided the plot of cylinder pressure Vs. cylinder volume on a log-log scale. Since both adiabatic compression and adiabatic expansion during power stroke are polytropic processes, straight lines would be observed for both these processes on log-log scale. The selection of two points on these two lines

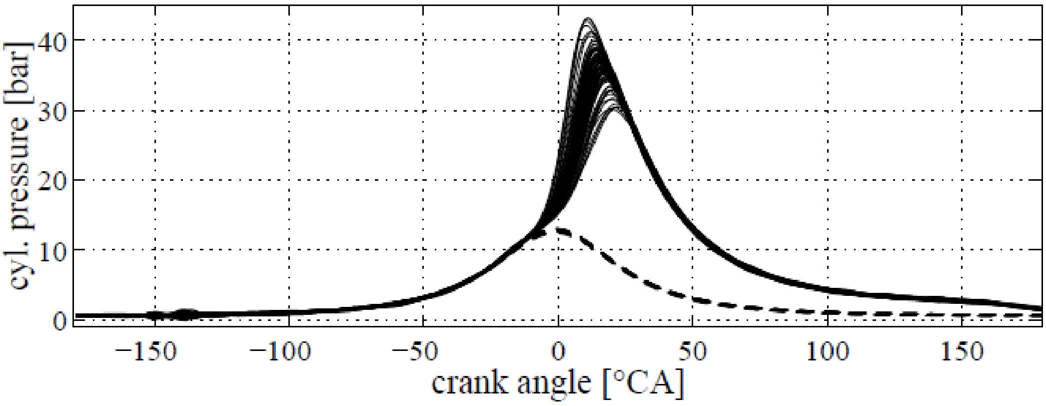


Figure 5.3 : Cylinder Pressure variation Curve

when piston is at same position and finding the pressure difference $\Delta p_c = p_{c2} - p_{c1}$. Guzzella L. (2004, pp: 167) indicated that this pressure difference varies almost linearly with the combustion energy and this pressure difference represent the average mechanical work on the piston by the igniting fuel. Using Δp_c, average piston speed and piston area, power acting on piston can be estimated. The estimation of cylinder pressure can be carried out using the technique proposed by Yaojung S and Moskwa J (1995, pp: 70-78). The estimated pressure can be plotted against volume to find Δp_c.
In this work the typical value of cylinder peak pressure is taken and taking crankshaft speed as input, acceleration signal is analyzed using observer to validate the input.

5.2.5 Model Parameter Estimation

The movement of piston exhibit a periodic behavior with same fundamental frequency as that of rotational speed of engine shaft. This provides a heuristic guideline to choose the value of k_3 (in Eq 5.16) as a function of crankshaft angular speed. The empirical choice is validated using simulation and experimental results reported later.

$$k_3 = \omega^2 = (2\pi N)^2 \quad \text{(Eq 5.22)}$$

---

where N is engine speed in revolution per second.

During experimental verification load is also applied by friction. Most frictional models described in literature are based on empirical relations as a polynomial in engine speed. A simplified frictional model is chosen with term containing only square of engine speed. The constant term representing the load acting on engine is also considered as a parameter whose value is defined as a polynomial in crankshaft speed as:

$$k_2 = b\,\omega^2 + c \qquad \text{(Eq 5.23)}$$

On the basis of simulation and experimental results it is established that the optimal selection of value of b varies between 0.02 and 0.5.

The value of parameters k_1 and k_2 depend on the operating point. The block diagram of switched linear system is shown in Figure 5.4.

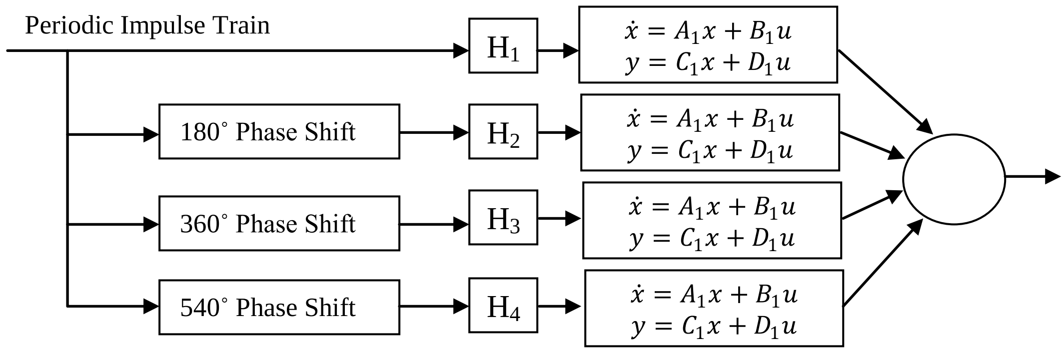


Figure 5.4: Block diagram for simple hybrid model of SI engine

5.2.6 Generation of Reference Signal for Fault Detection

The input to the model is assumed in the form of pulse. The updated model can be defined at any operating point. The linear model can be solved easily to find the system response representing the engine speed. This response can be used as a reference signal for comparison with the actual engine response to generate the residuals for the analysis of faults in system.

5.2.7 Results from Hybrid Model

An analysis of hybrid model indicates following results which are useful in statistical analysis of system.

1. Four peaks would be observed in one ignition cycle of a four cylinder SI engine. (Lemma 3)
2. Amplitude of four observed peaks represents four independent events. (Theorem I)
3. Crankshaft speed is proportional to input power. (Due to linear model of subsystems)
4. Crankshaft speed is proportional to amount of intake air. (By Corollary 3 and Proposition 2)

Having a good knowledge of system, the results can be also be visualized on heuristic grounds. However among the existing SI engine models, MVM can also be used to establish the third and fourth results, but cannot deduct the first two results. The results and properties of hybrid model will now be used to develop a statistical model for SI engine.

5.3 Statistical Analysis of Input to Hybrid System

Kami´nski T. *et al* (2003) studied the instabilities of combustions and random cycle to cycle variations of torque as nonlinearities and its control using spark advance as actuating parameter. Pulkrabek W. W. (1997, pp: 239) indicated that under ideal conditions, the combustion should be exactly same in all engine cylinders but this does not occur in practice due to variations that occur in the intake system. Even if no variations in intake occur, the turbulence within the cylinder would cause the statistical variations in engine output. The above studies indicate that randomness of the combustion process.

George V. [2006, pp: 341] mentioned that an integration of physical and stochastic modeling techniques, is useful to evaluate useful life as a function of uncertainties on account of certain faults. The main objective of statistical analysis is to study the statistical properties of peaks observed in the crankshaft speed fluctuation signal. For statistical analysis of crankshaft speed fluctuations, the random behavior of intake air would be considered first. The analysis would be carried out in the following steps:

- Determine probability density function (PDF) of random variable representing peak values of velocity observed in an ignition cycle.
- Formation of a collection of above random variable.
- Prove that above collection is Gaussian and Markov.
- Using corollary 4, the results are finally applied to crankshaft speed.

5.3.1 Determination of PDF of Peak Values of Crankshaft Speed

Using Result 4 (Section 5.2.7), it can be concluded that PDF of crankshaft velocity and air intake in engine cylinders are similar. The problem of finding the PDF of peak value of velocity is therefore reduced to find the PDF of intake air.

The PDF of air intake is estimated by a series of three hypothetical experiments representing the suction of air in engine cylinders. Hypothetical experiment is a statistical experiment which is not actually conducted but statistical properties of events generated by it could be analyzed.

5.3.1.1 Hypothetical Experiment 1

Consider a hypothetical experiment of counting the number of molecules sucked by cylinder as piston moves by a differential amount $\delta x \rightarrow 0$. The sample space for this hypothetical experiment would be $\{N_{min}, N_{min}+1, \ldots \ldots \ldots, N_{max}\}$, where N_{min} represents minimum number of air molecules that can be sucked during the differential movement δx and N_{max} is defined vice versa. Each differential movement δx of piston, where $\delta x \rightarrow 0$ during suction stroke produces an event of this experiment. A random variable ψ_i is defined on sample space that assigns some probability P(.) to each element of this sample space. The probability density function (PDF) for this variable can be approximated as uniform and IID (Independent but Identically Distributed) because amount of air sucked in cylinder depends upon the pressure difference between cylinders and manifold and under ideal conditions suction stroke of SI engine occur at constant pressure.

5.3.1.2 Hypothetical Experiment 2

The second hypothetical experiment is defined as counting the total number of molecules sucked by cylinder as piston moves from TDC to BDC. Each suction cycle would generate an event of this experiment. A random variable for events of this experiment can be expressed as a sum of large number of samples of Hypothetical Experiment 1:

$$\xi_k = \sum_i \psi_i \quad \text{(Eq 5.24)}$$

Using central value theorem, it can be concluded that PDF of this variable is Gaussian distribution.

5.3.1.3 Hypothetical Experiment 3

This hypothetical experiment is defined as counting the maximum number of molecules sucked by any of the four cylinders during an ignition cycle. A random variable of this experiment also represent sum of large number of samples of Experiment 1 and hence is a Gaussian variable. If $X_{m,i}$ is a random variable that maximum air is sucked in i^{th} cylinder in m^{th} ignition cycle where $i \in \{1, 2, 3, 4\}$ and $m \in \{1, 2, 3, 4, \dots \dots\}$ then $X_{m,i}$ is a Gaussian variable.

Defining a collection Z of $X_{m,i}$ and ignoring the index i for simplicity:

$$Z = \{X_1, X_2, \dots \dots \dots \dots \dots \dots X_n\}$$

Where n represents the number of sample and can be very large. The collection Z is our variable of interest which is claimed as Gaussian and Markov process.

Result 2 (Section 5.3.1.4) ensure the independence of events of collection Z. The method of proof is adopted from Speyer J. L. (2008, pp: 153-159) and applied to the problem at hand. In first step it would be proved that collection is Gaussian and then the collection would be proved to be Markov.

5.3.1.4 Collection is Gaussian

The characteristic function of collection is:

$$\Phi_Z(\omega_1, \omega_2 \dots . \omega_n) = E[e^{j\omega^T Z}] \quad \text{(Eq 5.25)}$$

where ω is the frequency variable. The exponent can be expanded as:

$$\omega^T Z = \omega_1 X_1 + \omega_2 X_2 + \cdots \dots . . + \omega_n X_n \quad \text{(Eq 5.26)}$$
$$\omega^T Z = \omega_n (X_n - X_{n-1}) + (\omega_n + \omega_{n-1})(X_{n-1} - X_{n-2}) + \cdots + (\omega_n + . . + \omega_1) X_1$$

Here $X_i - X_{i-1} = \Delta X_i$ represents the difference between peak values of air sucked during two successive ignition cycles. Eq 5.25 therefore becomes:

$$\Phi_Z(\omega_1, \omega_2 \dots . \omega_n) = E[e^{j\omega_n \Delta X_n} e^{j(\omega_n + \omega_{n-1})\Delta X_{n-1}} \dots e^{j(\omega_n + \omega_{n-1} + \dots + \omega_1)X_1}] \quad (\text{Eq } 5.27)$$

$$\Phi_Z(\omega_1, \omega_2 \dots . \omega_n) = \Phi_{\Delta X_n}(\omega_n)\Phi_{\Delta X_{n-1}}(\omega_n + \omega_{n-1}) \dots \Phi_{X_1}(\omega_n + \omega_{n-1} + .. + \omega_1)$$

The collection would be a Gaussian process if its characteristic function is Gaussian. As X_1 is Gaussian, the collection would be a Gaussian if ΔX_i is also Gaussian.

$$\Delta X_n = X_n - X_{n-1}$$

$$X_n = X_{n-1} + \Delta X_n$$

X_n and X_{n-1} represent maximum air sucked during different strokes of two different ignition cycles of engine. Using corollary 2 (Section 5.2.3.9), these strokes are independent so X_n and ΔX_{n-1} are independent. The characteristic function of X_n becomes:

$$\Phi_{X_n} = \Phi_{X_{n-1}}\Phi_{\Delta X_n} \quad (\text{Eq } 5.28)$$

$$\Phi_{\Delta X_n} = \frac{\Phi_{X_n}}{\Phi_{X_{n-1}}}$$

But as both X_n and X_{n-1} are Gaussian

$$\Phi_{X_n} = e^{-\frac{\omega^2 \sigma_n^2}{2}} \text{ and } \Phi_{X_{n-1}} = e^{-\frac{\omega^2 \sigma_{n-1}^2}{2}}$$

Hence

$$\Phi_{\Delta X_n} = e^{-\frac{\omega^2(\sigma_n^2 - \sigma_{n-1}^2)}{2}} \quad (\text{Eq } 5.29)$$

Which is the characteristic function of a Gaussian random variable with zero mean and variance $\sigma_n^2 - \sigma_{n-1}^2$. This indicates that difference between maximum air sucked observed during two consecutive ignition cycles is Gaussian. Consider a collection of non-overlapping increment Y. The collection represents difference between maximum air sucked in any of the four cylinder during two successive ignition cycles for n ignition cycles

$$Y = \{ X_1, X_2 - X_1, \dots \dots \dots \dots \dots \dots X_n - X_{n-1} \}$$ (Eq 5.30)
$$Y = \{ X_1, \Delta X_2, \dots \dots \dots \dots \dots \dots \Delta X_n \}$$

As per Section 5.2, suction in different cylinders is independent, which ensure that events of set Y also form set of independent events. The distribution function of non-overlapping increment of collection can be written as:

$$f_{X_1,..\Delta X_n}(x_1, x_2 - x_1, \dots, x_n - x_{n-1}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i - x_{i-1})^2}{2\sigma^2}}$$ (Eq 5.31)
$$f_{X_1,..\Delta X_n}(x_1, x_2 - x_1, \dots, x_n - x_{n-1}) = \prod_{i=1}^{n} f_{\Delta X_i}(x_i - x_{i-1})$$

where, x_0 is assumed to be 0. This indicates that Y is Gaussian. Using Eq 5.31, it is established that collection Z is Gaussian, and independent increment.

5.3.1.5 Collection is Markov

Consider by definition:

$$F_{X_n|X_1,,X_{n-1}}(x_n|x_1,.,x_{n-1}) = P(X_n \le x_n | X_1 = x_1,.,X_{n-1} = x_{n-1})$$ (Eq 5.32)

Given the past sequence, the right hand side of equation can be transformed in terms of increments.

$$F_{X_n|X_1,,X_{n-1}}(x_n|x_1,.,x_{n-1}) = P(X_n - X_{n-1} \le x_n - x_{n-1} | X_k - X_{k-1} = x_k - x_{k-1})$$ (Eq 5.33)

where k=1,2,….,n-1

The independent increment property of collection enables us to change the conditional probability with unconditional probability. Therefore

$$F_{X_n|X_1,,X_{n-1}}(x_n|x_1,.,x_{n-1}) = F_{\Delta X_n}(x_n - x_{n-1}) \qquad \text{(Eq 5.34)}$$

It has been proved earlier that ΔX_i is also Gaussian

$$F_{\Delta X_n}(x_n - x_{n-1}) = \int_{-\infty}^{x_n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(\eta - x_n)^2}{2\sigma^2}} d\eta \qquad \text{(Eq 5.35)}$$

$$F_{\Delta X_n}(x_n - x_{n-1}) = \int_{-\infty}^{x_n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\eta^2 - x_n^2 + 2\eta x_n}{2\sigma^2}} d\eta$$

$$F_{\Delta X_n}(x_n - x_{n-1}) = \frac{\int_{-\infty}^{x_n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\eta^2 - 2x_n^2 + 2\eta x_n}{2\sigma^2}} d\eta}{\frac{e^{-x_n^2}}{\sigma\sqrt{2\pi}}}$$

$$F_{\Delta X_n}(x_n - x_{n-1}) = \frac{\int_{-\infty}^{x_n} f_{X_n X_{n-1}}(\eta, x_{n-1}) d\eta}{f_{X_{n-1}}(x_{n-1})}$$

$$F_{\Delta X_n}(x_n - x_{n-1}) = F_{X_n|X_{n-1}}(x_n|x_{n-1})$$

Therefore Eq 5.34 becomes:

$$F_{X_n|X_1,,X_{n-1}}(x_n|x_1,.,x_{n-1}) = F_{X_n|X_{n-1}}(x_n|x_{n-1}) \qquad \text{(Eq 5.36)}$$

The collection Z therefore represents a Markov process. The events of this collection are the maximum amount of air sucked in any cylinder during an ignition cycle. By corollary 4, crankshaft speed is proportional to intake air. Hence collection of events generated by peaks observed in crankshaft speed is also Gaussian and Markov.

The basic philosophy of proposed fault diagnostic method is based on the peak velocities associated with four identical sub-systems. In a healthy engine, the largest peak observed in an ignition cycle can belong to any of the four cylinders with equal probability. Under faulty conditions, due to power loss, the smallest peak would correspond to faulty cylinder with highest frequency. The difference between two consecutive peaks is therefore taken as a measure of power loss due to faulty cylinder.

5.4 Summary

This chapter is aimed to the development of a hybrid model of SI engine. In this regard, a general description of hybrid systems is first provided and special emphasis is given to the properties of switched linear systems. A switched linear model for SI engine was then developed. The power generated inside the cylinders as a result of burning of air fuel mixture is considered as model input and crankshaft speed is considered as model output. Each of the cylinders is considered as individual sub-system of engine for which a transfer function is derived on the basis of basic laws of physics. The cylinder dynamics is considered as continuous sub-system. A discrete rule defining the periodic switching of ignition in different cylinders defined the discrete dynamics of engine model. The properties of hybrid model are studied. The hybrid model clearly identifies the properties that crankshaft speed would exhibit four peaks in one ignition cycle where the amplitude of four peaks depends on the input power and it was defined that the amplitude of four peaks represent four independent events.

The study of statistical events was based on the random variation of air sucked in each cylinder during an ignition cycle. It is established that the distribution of events representing the quantity of air sucked in cylinders during each ignition cycles is Gaussian. The properties of hybrid model and distribution of air input to engine cylinders is then used to establish that the random process representing random speed fluctuations observed on crankshaft in different ignition cycles is Gaussian and Markov.

Chapter 6

FAULT DIAGNOSTIC METHODOLOGY

The literature review indicates that Markov chain is a suitable tool for prediction, forecasting and estimation. The fault detection on the basis of Markov model is not studied widely for SI engine. This chapter presents the proposed misfire fault detection method based on Markov chain where the misfire is defined in Section 1.5. Before the development of method it would be appropriate to provide a brief overview of Markov processes and Markov chain.

The proposed methodology is developed under the inspiration of the fault detection methodology proposed by Rizzoni G. (1987 pp: 450-457) which has its roots in the model defining torque as:

$$T_i(t) = P_i(t).\, g(\theta) \qquad \text{(Eq 6.1)}$$

The details of method proposed by Rizzoni G. are presented in section 3.4.2. It is established that crankshaft speed is depend on the cylinder pressure and all the peaks of torque/ speed are processed as scalar quantities in the form of a single array.

Rizzoni G. (1987 pp: 450-457) used the correlation analysis to isolate the fault. In his analysis, Rizzoni G. collected the signal patterns of all possible engine misfire faults on the operating engine speed. The observed data is correlated with all those signals to identify the similarity of the observed signal with possible fault pattern. The basic advantage of the proposed fault diagnostic methodology based on hybrid model over the method proposed by Rizzoni G. is that the latter method detects the fault by maximum likelihood method and used additional steps of correlation analysis to isolate the fault. The proposed method on the other hand can detect and isolate the faulty cylinder in a single algorithmic step.

The proposed method is based on hybrid model, with number of sub-systems equal to the number of cylinders in the system. This instigated to de-multiplex the measured speed into four streams for a four cylinder engine so that each stream can be associated to one of the sub-systems. For data de-multiplexing, SI engine provides features that can not only be used to distribute the inputs into four streams but also help to synchronize the four streams with the active cylinders. The details of fault detection methodology are described in section 6.2.

6.1 Markov Chain

Papoulis A. [2006, pp: 695] defined Markov processes as the simplest generalization of independent processes in which outcome at any instant depends on the outcome of preceding event only i.e. for Markov process past has no influence on the future if present is given. A large number of random processes like gambler ruin problem, random walk and branching processes etc are being modeled using Markov chains.

6.1.1 Properties of Markov Chains

The terminology and properties of Markov chains can be observed in any text book on stochastic process. In this section only those properties of Markov chains would be highlighted that are pertinent to the work presented in this thesis.

6.1.1.1 Markov Assumption

The basic property of Markov process is the Markov assumption i.e. the future evolution of states $x(t_n)$, of the process is independent of the past record of process states (i.e. before time t_n), given a complete description of the current states of process at time t_n. The Markov process can mathematically be described as:

$$p[x(t_n) \leq x_n / x(t), t \leq t_n] = p[x(t_n) \leq x_n / x(t_{n-1})]$$
$$where\ t_1 < t_2 < \cdots < t_n \qquad \text{(Eq 6.2)}$$

Where p is the probability.

6.1.1.2 Definition of States

A random process X is a Markov chain if it satisfies the Markov assumption given in Eq 6.2. where all the possible states $x_1, x_2, \ldots, x_n \in S$, where S is a countable set.

6.1.1.3 State Transition

The property of Markov assumption ensures that next state of system can be estimated using the information of only current state of system. A state transition matrix defines the probabilities of change of system states from one state to any other state. In any stochastic process, the state transition matrix can be defined by an analysis of probabilities of jump between different states of system.

In a state transition matrix P, each row represents the probability of jump from that state to some other state e.g. p_{11} represents the probability of jump from state 1 to itself. Similarly p_{12} represent probability of jump from state 1 to state 2. The matrix P is a stochastic matrix and exhibits the following properties:

- All the entries of matrix P are non-negative i.e. $p_{ij} \geq 0 \; for\ every\ i\ and\ j$
- The sum of all entries of each row of matrix P is unity i.e. $\sum_j p_{ij} = 1 \; for\ all\ i$

Given the current state of system, matrix P can be used to predict the next state of the system as:

$$p(1) = p(0)P \qquad \text{(Eq 6.3)}$$

where $p(0)$ is the initial probability and $p(1)$ is the predicted probability after single steps. The evolution of states is dependent on the occurrence of discrete events.

6.1.1.4 State Prediction

Given the initial state of system and state transition matrix, the state of system at any instant can be estimated. The n-step transition probabilities can be estimated using Chapman-Kolmogorov equation defined as:

$$p(n) = p(0)P^n \qquad \text{(Eq 6.3)}$$

where $p(0)$ is the initial probability and $p(n)$ is the predicted probability after n steps. The evaluation of n^{th} power of state transition matrix is performed by eigenvalue decomposition of matrix P as:

$$P = U\Lambda U^{-1}$$

$$then \qquad P^n = U\,\Lambda^n U^{-1} \qquad \text{(Eq 6.4)}$$

6.1.1.5 Stationary Distribution

A probability vector $\pi = [x_1\ x_2\ ...\ x_n]^t$ is said to be the stationary distribution of a finite Markov chain if

$$\pi = \pi P \qquad \text{(Eq 6.5)}$$

i.e. the probability would not improve any further. In any irreducible Markov chain having finite number of states, at least one stationary distribution exists. The stationary distribution represents the limiting state probability of the system.

In any system the states of the system can be classified into a number of communicating classes. A communicating class C is a set of states such that if $i \in C$ then $j \in C$ if and only if $i\ and\ j$ communicate.

The basic problems of developing a diagnostic algorithm based on Markov chain are :

- To define a stochastic process that can be represented as a collection of state.
- To ensure that the proposed stochastic process is Markov
- To ensure that the fault could be detected by the probabilistic analysis of states of system

Hybrid model presented in chapter 5, however provide sufficient physical insight to define the desired stochastic process.

6.2 Fault Diagnostic Method

The first step in the development of fault diagnostic methodology is the selection of appropriate variable that satisfies the conditions of Markov process. The variable would be transformed into a finite set of states. An event would also be defined for the evolution of states. The defined states would then be used for residual generation.

The residual signal is evaluated using Markov chain by analyzing the limiting probability of states, assuming a healthy current state. A threshold can be selected on the components of the limiting probability vector to identify the fault. Each step of fault diagnostic methodology is presented in this section.

6.2.1 Selection of Random Variable

The proposed fault diagnostic method used the output of hybrid model i.e. crankshaft speed for the generation of residual. The hybrid model and its properties are presented in chapter 5. Following are the main properties of interest for the development of fault diagnostic algorithm.

- Four peaks are present in the plot of velocity waveform of SI engine during an ignition cycle.
- The four peaks of crankshaft speed can be indexed by observing the cylinder number which is active when the peak is observed.
- The event of observing largest peak during an ignition cycle is a random variable.
- The collection of events representing largest peak form a Markov process.

On the basis of mentioned model properties, the four peaks observed in the crankshaft speed signal is taken as the variable of interest which is Markov. It is assumed that the data acquisition rate would be large enough to observe those peaks.

6.2.2 Selection of Event

One set of peak values would be taken for each ignition cycle. The ignition cycle is therefore considered as the desired index to update the probabilities.

6.2.3 Residual Generation

In the proposed fault diagnostic technique, a reference signal would be generated by first tuning the model with the actual engine so that the output (crankshaft speed) of model matches with the actual engine. The difference between the four peaks estimated through model and observed through engine would be used as a residual as shown in Figure 6.1. Under ideal conditions, the model would be running run time in parallel for the generation of residual.

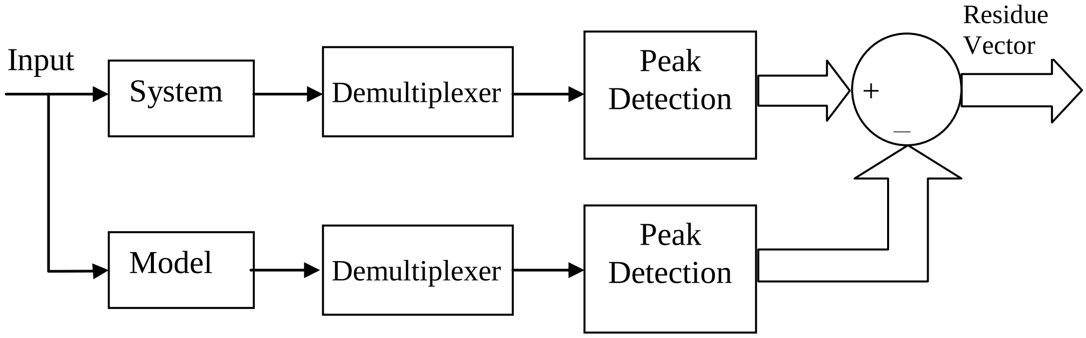


Figure 6.1 : Residual Generator with model running in Parallel

If ω^T represent the vector indicating the four peaks of AC component of crankshaft speed signal within an ignition cycle as computed through model (i.e. a reference signal) then:

$$\widehat{\omega}^T \triangleq [\omega_1 \quad \omega_2 \quad \omega_3 \quad \omega_4] \tag{Eq 6.6}$$

In practice fault diagnosis would be carried out under steady state condition at constant engine speed. Lemma 3, given in section 5.2.3.4 indicated the property of hybrid model that for a healthy and ideally balanced engine operating at constant speed *all the four peaks of crankshaft speed fluctuation signal during an ignition cycle would be identical*. Hence the model based computation would result in four identical peaks. Therefore the peak value would become a constant and need not be estimated run time but for defined operating conditions, the value can be estimated offline by first tuning the model with the healthy engine. The estimated value can then be used to generate the reference signal.

If $\|\omega\|$ represent the values of four peaks estimated through model, then the reference vector $\widehat{\omega}^T$ will become:

$$\widehat{\omega}^T \triangleq \|\omega\|.[1 \quad 1 \quad 1 \quad 1] \tag{Eq 6.7}$$

The residual generation scheme is therefore simplified significantly.

In the simplified approach, instead of running the model in parallel, the estimated value of vector $\widehat{\omega}^T$ would be fed to the residual generator for on-line calculations. The block diagram of new residual generation scheme is shown in Figure 6.3

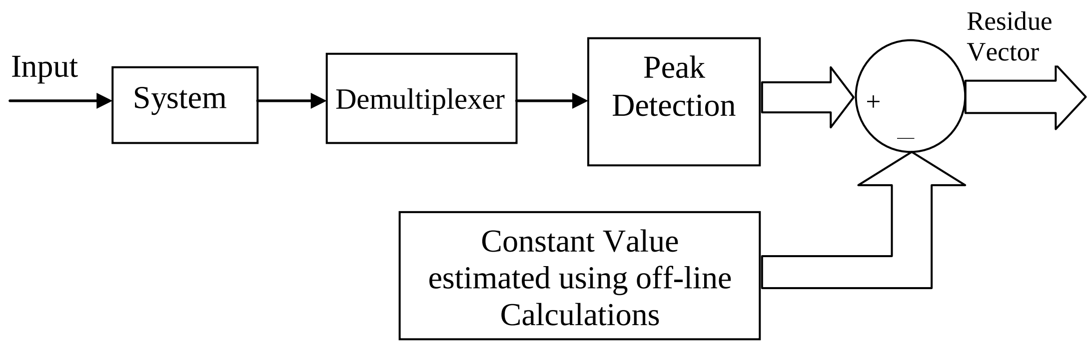


Figure 6.2: Simplified Residual Generator with off-line reference estimation

Using reference vector signal $\widehat{\omega}^{\,T}$ and the vector ω^T representing the peak values of data during an ignition cycle, residual vector can be defined as:

$$r^T(k) = \widehat{\omega}^{\,T}(k) - \omega^T(k) \quad \text{(Eq 6.8)}$$

Here each component of residual vector represents the variation of each cylinder from a reference behavior. The proposed residual scheme satisfies the property of residual mentioned in Section 2.5.1 that value of residual is zero under no fault condition.

6.2.4 Residual Evaluation

In Chapter 5 it was already proved that the process defined by observing the peak value of crankshaft speed during an ignition cycle forms a Markov process. It can therefore heuristically be concluded to use Markov Chains for *Residual Analysis*. Markov Chains have a potential to identify a small biasing condition by analyzing the limiting probability.

For application of Markov chain it is necessary to define the states. Since residual vector is generated once in one complete ignition cycle, the states would also be updated once after complete ignition cycle.

The state of Markov chain can be defined heuristically with physical significance behind it as *the index of cylinder showing the maximum non-uniformity or power loss*. This corresponds to the row number corresponding to the largest value of component in the residual vector. In this way a set of four possible states can be defined for engine. During engine operation, a state would be assigned for each ignition cycle and the states would jump from one state to the other state in successive ignition cycles.

The set of four states $s_i \quad i = 1,2,3,4$ are therefore defined as:

$$
\begin{aligned}
&s_1\text{: index}(\ \|d_i\|_\infty) = 1;\\
&s_2\text{: index}(\ \|d_i\|_\infty) = 2;\\
&s_3\text{: index}(\ \|d_i\|_\infty) = 3;\\
&s_4\text{: index}(\ \|d_i\|_\infty) = 4;
\end{aligned}
\qquad \text{(Eq 6.9)}
$$

Where $\|d_i\|_\infty$ represent the infinity norm of vector d_i (i.e. the largest element of vector d_i) and index($\|d_i\|_\infty$) represents the index of the largest element of vector d_i. The presence of state is the indication, that the particular cylinder is assumed as faulty in that specific ignition cycle. The decision of fault would however not be made on the basis of single ignition cycle but the data of a large number of ignition cycles would be complied to decide about the fault in step of residual evaluation.

For fault analysis a state transition matrix F is defined representing frequency of transition from one state to some other state. The matrix F can be generated by observing the peak values of speed in the observed data set. We define a single state transition event from state s_i to s_j in m^{th} ignition cycles as a matrix with 1 in j^{th} row and i^{th} column and 0 elsewhere e.g.

$$
F_m = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(Eq 6.10)}
$$

Represents that in mth ignition cycle maximum power loss is observed in 3rd stroke and in (m+1)th ignition cycle maximum power loss is observed in 2nd stroke.

Again defining a matrix F as:

$$F = \sum_{m} F_{m} \qquad \text{(Eq 6.11)}$$

The matrix F contains the frequency of occurrence of all state transitions as:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ f_{31} & f_{32} & f_{33} & f_{34} \\ f_{41} & f_{42} & f_{43} & f_{44} \end{bmatrix} \qquad \text{(Eq 6.12)}$$

where, f_{ij} represents the frequency of arrival of i^{th} state from j^{th} state. The total number of arrival to i^{th} state from any other state is the sum of i^{th} row i.e.

$$f_i = \sum_{j=1}^{4} f_{ij}\ ,\ i = 1,2,3,4 \qquad \text{(Eq 6.13)}$$

Matrix F is then converted to a probability transition matrix P. The elements p_{ij} of matrix P represent the probability of state transition from ith state to jth state, where i and j belong to set {1,2,3,4}. For a four cylinder engine the dimension of matrix P is 4×4. The matrix *F* would then be used to calculate the transition probability matrix P that satisfies the condition of transition probability matrix of a Markov chain:

$$\begin{aligned} &p_{ij} \geq 0; \qquad i = 1, 2, 3, 4 \\ &\textstyle\sum_{j=1}^{4} p_{ij} = 1;\ \ i = 1, 2, 3, 4 \end{aligned} \qquad \text{(Eq 6.14)}$$

The transition probability matrix P is obtained by dividing all the elements of i^{th}row of matrix F with the frequency of arrival i^{th} state i.e.

$$P = \begin{bmatrix} \frac{f_{11}}{f_1} & \frac{f_{12}}{f_1} & \frac{f_{13}}{f_1} & \frac{f_{14}}{f_1} \\ \frac{f_{21}}{f_2} & \frac{f_{22}}{f_2} & \frac{f_{23}}{f_2} & \frac{f_{24}}{f_2} \\ \frac{f_{31}}{f_3} & \frac{f_{32}}{f_3} & \frac{f_{33}}{f_3} & \frac{f_{34}}{f_3} \\ \frac{f_{41}}{f_4} & \frac{f_{42}}{f_4} & \frac{f_{43}}{f_4} & \frac{f_{44}}{f_4} \end{bmatrix} \qquad \text{(Eq 6.15)}$$

When the cylinders of engine are maximally balanced and input to all cylinders is also exactly same, the current state may jump to any of the four possible states with equal probability and hence the probability of all possible jumps is 0.25. It is therefore anticipated that under no fault condition all the entries of matrix P would be close to 0.25.

Using transitional probability matrix P and a vector $p(0)$ that define the initial fault probability of four cylinders, Fault probability after n transitions is given by:

$$p(n) = p(0)P^{(n)} \qquad \text{(Eq 6.16)}$$

Using eigenvalue decomposition, the above expression can be written as:

$$p(n) = p(0)VD^{(n)}V^{-1} \qquad \text{(Eq 6.17)}$$

Where V is a matrix of eigenvectors and D is a diagonal matrix with eigenvalues of transition probability matrix on diagonal. As $n \to \infty$, the limiting state probability would be calculated. Since matrix D is diagonal, the calculation of arbitrary power of matrix is simply a computation of scalar power.

The matrix F is formed by checking the peak values of an ignition cycle at run time and finding the state transitions. Having a batch of data of 100 or more ignition cycles, probability transition matrix is calculated. Assuming initial fault probability vector *p(0)=[0.25 0.25 0.25 0.25]* limiting probability can be found offline.

Since the step of evaluation of eigenvalue decomposition of 4×4 matrix is not very easy step to be programmed in conventional microcontrollers, the method is suitable for off- line analysis.

6.2.5 Threshold Definition

The elements of vector p represent the probability of fault in each engine cylinder. Brotherton T et-al, [2000, pp: 163-171], mentioned that for nominal operation, the residuals are represented by white noise with small variance and in the presence of fault the variance would be increased. It can therefore be anticipated that when all the cylinders are maximally balanced, the components of limiting probability vector would remain same and remain close to 0.25. Under faulty condition the value component corresponding to the faulty cylinder would become largest but remain less than 1. The value of threshold can be chosen between 0.25 and 1. The experimental results given in next section indicates that limiting probabilities converge to the cylinder exhibiting largest non-uniformity.

6.2.6 Comparison of Method with Approach of Rizzoni

The presented approach of residual generation is similar to the residual approach adopted by Rizzoni G (1987 pp: 450-457). The main difference between the approach adopted by Rizzoni and that proposed in this thesis is that Rizzoni used a complete data based approach in which the magnitude of the reference signal was defined by observing the L_1 norm of the complete data but in this approach, magnitude of reference signal is defined by the proposed hybrid model. In the approach of Rizzoni, if the amplitude of any specific peak was increased due to some disturbance or noise then that peak if represented L_1 norm would be taken as the reference signal although it is not representing the actual engine speed.

The theory developed using hybrid model is however coherent with the Rizzoni approach. If we isolate the link of constant value in Figure 6.2 from the model and take any arbitrary number (which is larger than the number actually estimated through model) to generate the residual then residual analysis would still provide the correct fault detection and isolation.

Rizzoni G. (1988, pp: 237-244) used the crankshaft speed as a scalar and defined the residual as non-uniformity in speed. A plot of non-uniformity histograms was provided by Rizzoni which indicates that an increase in non-uniformity of crankshaft speed would shift the mean value of histograms from zero to larger values. By generating residuals using this method, the misfiring cylinder could however not be identified directly. For fault isolation Rizzoni G. used the method of autocorrelation described in section 3.3.2, in which the observed signal was compared with signal samples taken under different fault conditions.

In the proposed method histogram can be plotted for each component of residual vector, where each component corresponds to a cylinder. The plot of histograms of four components of residual vector indicates that the histograms of each component of residual vector overlap under no fault condition and the histograms of each individual component shift from each other in the event of fault. Using relative shift of each vector component, misfiring can be identified directly using the residual. A simulation of outputs (Engine speed) of each sub-system both under misfire and no misfire condition is provided in chapter 7.

The shifting of histograms can be used for developing a residual evaluation method for detecting the misfire condition. This shifting of histograms depends upon:

- Non-uniformity between individual cylinders (Misfire conditions)
- Engine operating speed (Current condition)
- Load on engine

The presented method can be considered as simple generalization of method presented by Rizzoni.

The presented method share same limitations as the method proposed by Rizzoni i.e. at higher engine speed and smaller load condition, the shifting of histograms would be smaller. The value of residual vector depends upon the operating conditions of engine.

The proposed hybrid model however provides few more simplifications for the generation of residual. The properties of hybrid model indicate that the largest speed drop is observed when no pressure would be generated inside an engine cylinder i.e. when a misfire event occurs in engine cylinder. If residual is generated using this approach, even the reference signal would not be required and the algorithm proposed on model based analysis could be transformed to a data based algorithm.

In this residual generation scheme the variation of peaks observed between two consecutive ignition strokes would be observed. This residual scheme has physical significance that how much power is dropped between two consecutive power strokes. In this case the residual vector can be written as:

$$
\begin{aligned}
&r(k) = [\omega_1 - \omega_2 \quad \omega_2 - \omega_3 \quad \omega_3 - \omega_4 \quad \omega_4 - \omega_1] \\
&or \\
&r(k) = [d_1 \quad d_2 \quad d_3 \quad d_4]
\end{aligned}
\qquad \text{(Eq 6.18)}
$$

The proposed residual scheme also satisfies the property of residual mentioned in section 2.5.1 that value of residual is zero under no fault condition. For ideally balance engine, under no misfire condition, all the peaks of crankshaft velocity would be equal and the difference would always be zero.

6.2.7 Data Based Approach

The major finding of result is however that the indices of largest diagonal element in F matrix is same as that of faulty cylinder, hence fault can also be estimated using F matrix only. This not only saves the computational load of formation of matrix P and SVD but also enables the algorithm to detect the fault on-line. The simplicity of algorithm makes its implementation possible on ordinary microcontrollers.

The Figure 6.3 presents a general flowchart of the implementation of algorithm on a microcontroller. It is known that the gear has 12 regular teeth at an angular displacement of $1/12^{th}$ of a revolution and an additional tooth between two regular teeth as a double tooth. During each stroke of an ignition cycle, six teeth would appear and in this way instantaneous angular speed of crankshaft can be defined at six points during an ignition cycle.

The basic implementation philosophy is that based on the measurement of time taken to traverse an angular displacement of 30°. Each low to high or high to low transition of gear tooth is an indication of completion of this angular displacement except for the case of double tooth that must have to be dealt separately. In the algorithm proposed in this thesis, we are not interested in the exact value of crankshaft speed but only we

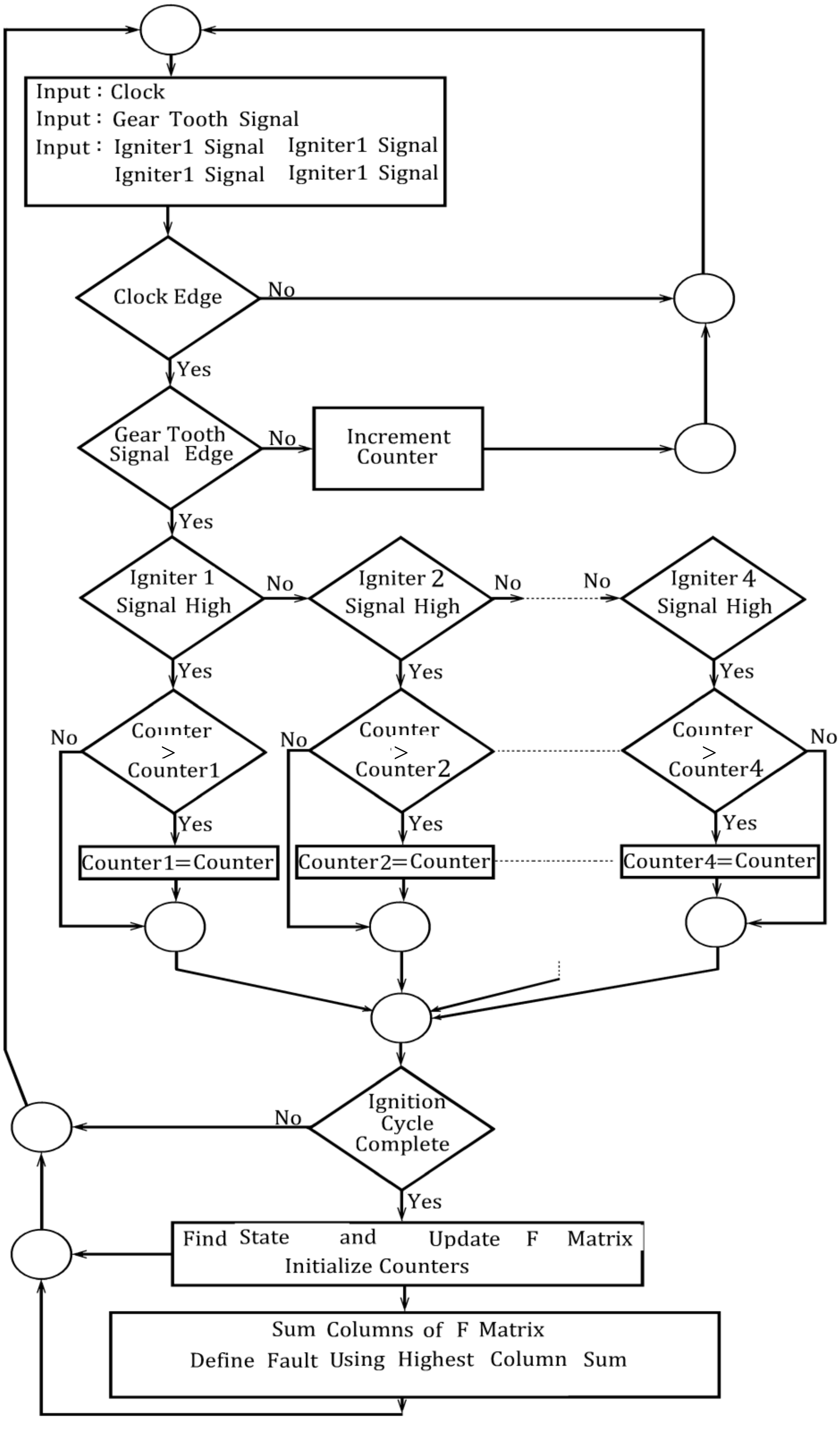


Figure 6.3: A Data based algorithm of Misfire detection for implementation in microcontroller

want to know whether crankshaft speed become low during ignition cycle of any specific cylinder with higher probability. Ignoring the double teeth, we can get six readings of instantaneous crankshaft speed during each ignition stroke. If a timing signal is generated at sufficiently high rate then timing information between the occurrences of two consecutive teeth would be an indication of instantaneous crankshaft speed. The largest value of time would imply the smallest crankshaft speed. The smallest observed value of time corresponding to the ignition stroke of each cylinder is recorded. After the completion of each stroke the observed largest peak value of time corresponding to the peak value of crankshaft speed during that stroke. These largest time values would be recorded during all four strokes of an ignition cycle to define the F matrix. This information can be used to identify the misfiring cylinder as the smallest value of crankshaft speed occur during activity of misfiring cylinder with highest probability. The identity of cylinder in which power stroke is occurring can be obtained by observing the igniter signal .

The timing information can be generated by starting an interrupt signal. This interrupt signal will always increment a counter and hence content of counter is an indirect indication of timing. The counter would be incremented by interrupt signal till the receipt of edge signal of gear tooth. Larger is the value of counter longer is the duration of received pulses and smaller is the crankshaft speed. The identity of cylinder is established by the igniter signal and if the recorded time is larger than the existing value of observed largest time then new value of time is saved and otherwise the value is discarded.

Therefore during each stroke of cylinder largest value of time taken to sweep the given displacement of 30° is recorded from among the six possible readings. This value would correspond to the smallest instantaneous value of crankshaft speed observed during the stroke. This value would be recorded for the ignition stroke of each cylinder once during the ignition cycle.

After the completion of complete ignition cycle, the four peak times (smallest speed) would be present in four different variables in microcontroller. These peaks could then be used to form F matrix in the memory of microcontroller. After logging the data for 400 to 500 ignition cycles, fault could be identified.

6.3 Summary

This chapter is aimed to describe an algorithm to detect misfire fault in SI engine using Markov Chains. The chapter first presented the basic properties of Markov processes and Markov chains. It is identified that Markov chains can be used to predict the probability of its states.

It is already proved in Chapter 5 that the process representing the crankshaft speed fluctuations of an SI engine is a Markov process. In this chapter the process is first converted to a discrete state vector. These state vectors represent residuals for fault detection method. A method is formed to calculate the state transition probability matrix for the Markov Chains. It was claimed that useful information about the engine fault can be obtained by finding the limiting probability of different states. To evaluate the arbitrary power of state transition matrix, the eigenvalue decomposition can be used.

In the end it is also claimed that the results of fault diagnostic algorithm can also be obtained heuristically by the inspection of some intermediate matrices also. When the results are finalized on the basis of intermediate matrices, a very simplified data based algorithm is formed that can even be implemented in a microcontroller.

Chapter 7

RESULTS AND DISCUSSIONS

This chapter provides a brief description of development of experimental setup and its limitations, software tools used for data acquisition and analysis. The experimental setup and results of data analysis are also provided in this chapter. The data analysis is divided into two groups:

- Data analysis using simulated data
- Data analysis using experimental data

The experimental validation of presented results can also be divided into two major groups.

- Validation of Model
- Validation of Fault Detection Method

The validation of model was carried out by comparing the outputs of model simulation with the experimental data. The fault diagnostic algorithm was validated by introducing misfire fault in engine and detecting the fault through proposed fault diagnostic algorithm using experimental data. The scope of future research using the proposed methods is discussed in the end.

7.1 Experimental Setup

The experimental setup was also developed during the research work. The funds for the development of experimental setup were provided by ICT R&D Funds Pakistan. The ICT R&D funds were utilized for the purchase of a half cut car of make Honda. The setup was however not research friendly. Some necessary alterations were performed in purchased setup to facilitate experimental work. These alterations include:

- Fabrication of frame/ stand for setup
- Installation of additional sensors in system for necessary data acquisition

- Marshalling of signals from all sensors to a junction point to ensure easy access of sensor signals
- Experimental space was provided in setup where a data logging system could be placed during experiments
- Provision of hardware and software for data logging

The sensor used for the proposed misfire detection algorithm was already present in the vehicle. However to extend the application of setup for other research activities, some additional sensors were also installed in setup. The pictures of experimental setup, introduction of faults in system and signals acquired during experiment are provided in appendix A.

The setup offered two possibilities for data acquisition on computer.

- The data acquisition cards from National Instrument could directly be connected with vehicle sensors. The data could be acquired using “Labview” or “Signal Express”. This setup is shown in pictures provided in appendix A.
- OBD-II connecter was also available in the setup so that a fault diagnostic toolkit could be connected to the vehicle to read the system variables using a PC based OBD-II diagnostic software.

The basic weaknesses of experimental setup were:

- Dynamo was not installed in the setup and hence controlled value of load could not be applied on engine. The load could however be applied using either Head Lights, Air Conditioner or by allowing the wheels to rotate and applying brakes but the estimate of applied load was not available.
- In-cylinder pressure could not be measured due to non-availability of its sensor in the setup.
- During the experiments the throttle position was manually controlled by pressing the accelerator and load was applied by the application of brakes. The manual adjustment of these parameters also resulted in introduction of errors in results.

In model validation, an approximate value of in-cylinder pressure was chosen based on the information available in literature and input was tuned to match the simulation results with experimental results. The tuned values were then validated by comparing the response of model and experimental results with misfire fault introduced in system and keeping all other conditions unchanged.

7.2 Model Validation

The hybrid model proposed in section 5.2 was simulated using Matlab/ Simulink. All the sub-systems were receiving input from a periodic pulse train but the input pulses of each sub-system was phase shifted from each other. The period of input periodic pulse train was defined by the engine speed and the phase delay between inputs to different cylinders was equal to the one fourth of period of input pulse train. This ensured that all the cylinders were getting input in the cyclic manner defined by the hybrid model. Only steady state model was studied.

Since power generated by the burning gases depends upon the pressure inside the cylinder, the amplitude of input pulses was adjusted according to the peak pressure developed in cylinder. The gain of cylinder model was adjusted according to the formula provided in linear models derived in section 5.2, however the values of gain vary slightly.

To control the misfire events an additional gain element was added with each subsystem. This gain of all elements was given a value equal to 1 for no-misfire simulation. To simulate the misfire situation, the gain of the corresponding sub-system was set to zero so that its output did not participate in the net system output. The nominal values of model parameters/ constants used in simulation are provided in Table 7.1. Under no misfire condition, the model was tuned to match its output with the experimental results.

TABLE 7.1 PARAMETER VALUES USED IN SIMULATION

Parameter	Value	Description
m	20 Kg	Mass of Engine moving assembly
b	0.2	Friction Coefficient
k_3	10000	Elasticity Coefficient
γ	1.4	C_p / C_v
P	10 hp	Power generated in cylinder
η	0.3	Efficiency
ω	100 rad/s	Engine operating speed

Using same parameter values, the misfire situation was simulated. The simulation results of hybrid model are shown in Figure 7.1.

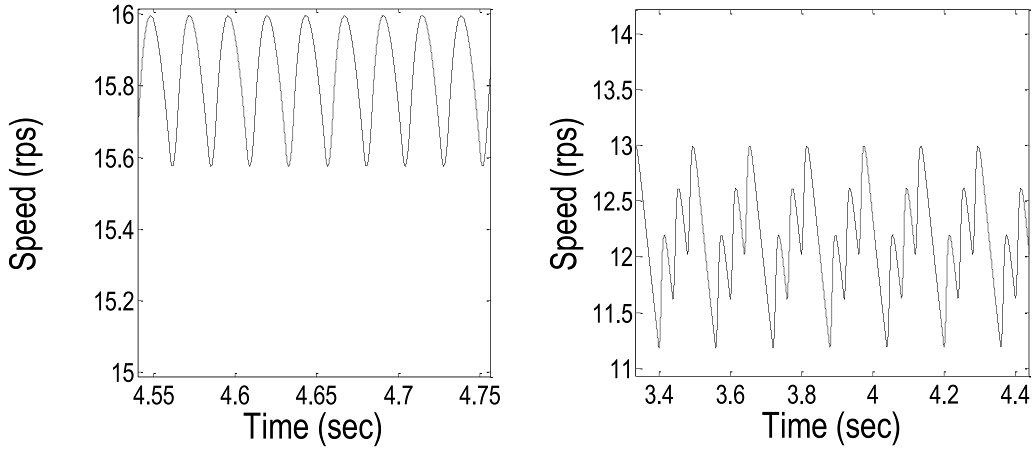


Fig. 7.1 Simulation Results: The waveforms representing fully balanced engine operation (left) one cylinder misfiring (right)

The simulation results were then validated by conducting an experiment. In the experiment the pulses generated by the magnetic sensor when a gear tooth come close to it were logged using a data acquisition card from National Instrument Inc. on an analog channel. A conceptual diagram of experimental arrangement of setup is shown in Figure 7.2 and actual experimental setup is shown in picture in appendix A.

The acquired signal was converted to pulses by comparing its signal level with a threshold value. The signal was polled at a constant rate using data acquisition cards. Following data is applicable to the experiment.

$$Number\ of\ Teeth\ in\ gear = 13$$

$$Angular\ spacing\ between\ normal\ Teeth = 30°$$

$$Angular\ spacing\ between\ double\ Teeth = 15°$$

$$Reference\ indication\ by\ Double\ teeth$$

$$Data\ Acquisition\ Rate = 50000\ samples/second$$

The reference was first searched by finding the double teeth. The number of samples polled in the time interval of passing of two consecutive gear teeth in front of magnetic sensor was observed. The number of samples polled was converted to time as:

$$Time = \frac{Number\ of\ samples\ polled}{Data\ Acquisition\ Rate}$$

Using angular displacement between two consecutive teeth and time to traverse that angular displacement, crankshaft speed was estimated. Crankshaft speed was finally

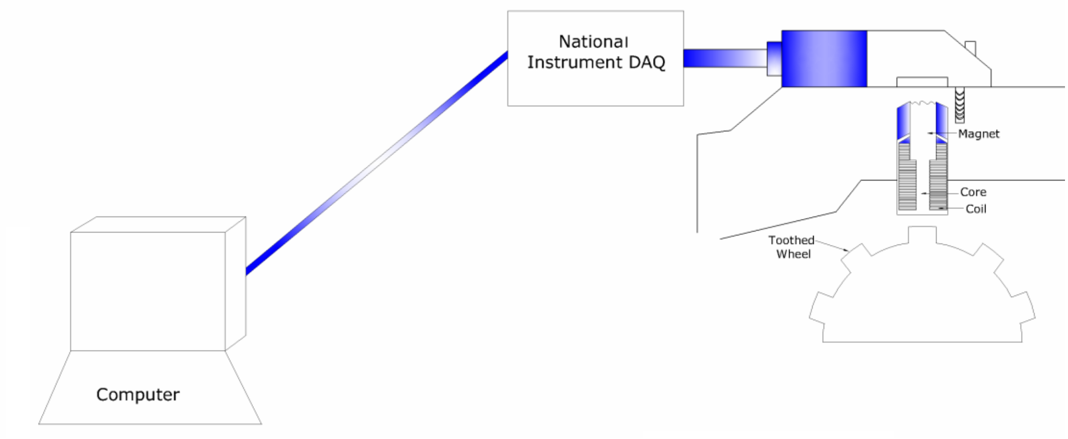


Figure 7.2 : Conceptual Diagram of Experimental Setup

plotted as a function of time. The experiment for the measurement of speed was conducted both under no-misfire condition and misfire condition. During experiment some load was kept on engine by application of brake. The value of applied load was however unknown but an effort was made to keep load similar in both experiments by retaining the brake paddle at the same position during both experiments. The experimental results are shown in Figure 7.3.

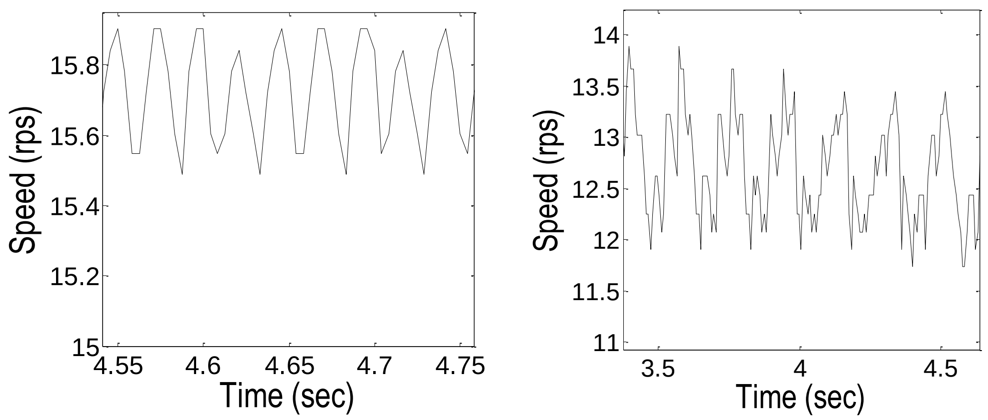


Figure 7.3 Experimental Results: The waveforms representing fully balanced engine operation (left) one cylinder misfiring (right)

During experiments misfire situation was introduced in engine by removing one spark circuit. The removed spark circuit is shown in picture in appendix A. A comparison of simulation and experimental results, shown in Figure 7.1 and Figure 7.3, indicates sufficient similarity to validate the model at the mentioned operating points.

7.3 Fault Detection Algorithm Validation

The experimental / simulation results for the proposed method of fault diagnostic are discussed in this section:

7.3.1 Fault Detection Method

The conventional models use crankshaft speed as a continuous variable for analysis. Hybrid model with its property of independence of four sub-systems, however provide enough justification to de-multiplex the continuous data stream into four independent data streams that can be analyzed independently. Having a data of two ignition cycles i.e. angular displacement of 8π, the first data stream would contain data from angular displacement of range 0 to π and data from 4π to 5π. Similarly the second stream would contain the data from π to 2π and data from 5π to 6π. The splitting of data and then recombining it with another data segment may result in non-smooth data for which continuous time analysis would become difficult. Under steady state engine operation, the problem is however not that significant as apparent in experimental data shown in Figure 7.3 (left) in which no significant jump at the joining point occurred when first pulse is joined with the fifth pulse. In the proposed data based methodology, even the occurrence of such a jump would not significantly affect the results.

Simulation is performed to study the behavior of proposed algorithm with or without misfire events. For simulations data was generated using proposed hybrid model. Since hybrid model is highly deterministic, noise equal to 10% of the observed AC component of signal was added in it to simulate the physical data.

7.3.1.1 Residual Generation

To evaluate the effectiveness of residual proposed in section 6.2.3 the engine output was simulated using hybrid model. The simulation was performed using fixed time interval. The period of input signal was selected according to the selection of engine speed i.e. half of engine speed in revolution per second (*rps*). Using input signal and fixed simulation time interval, the output was divided into four streams of data one corresponding to each subsystem. The problem of dividing the experimental data into four streams is discussed in section of experimental verification. Histograms of all the streams were plotted as in Rizvi [2009, pp: 93-100] and are shown in Figure 7.4.

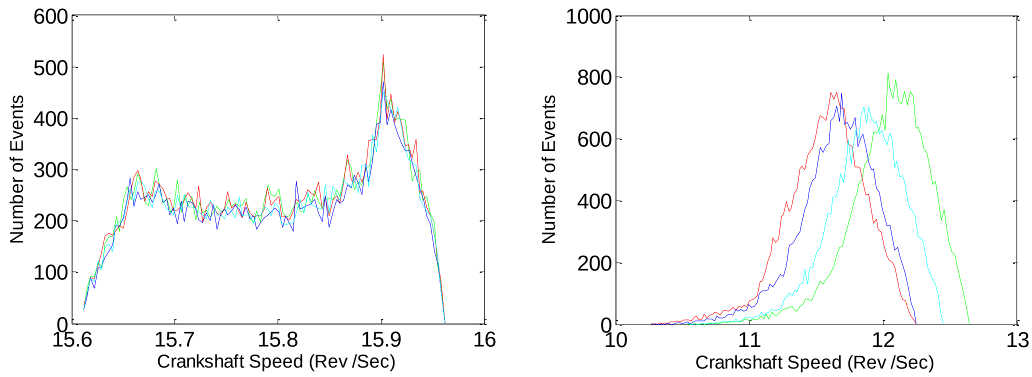


Figure 7.4 : Distribution of crankshaft speed during the activity period of four cylinders

Following results can be deducted by analysis of above simulation results:

- The speed fluctuation is fairly small under no fault condition while the speed fluctuations are fairly large under fault conditions. The basic reason for larger speed fluctuations are the biasing provided by fault to one of the sub-system that resulted in instantaneous drop in engine speed.
- The states defined in section 6.2.4 will occur with equal probability under no misfire condition, however under fault conditions, the probability of occurrence of state is defined by the fault intensity.
- The partial overlap of histograms associated with sub-systems indicates that residual alone if used for deciding fault would result in a large number of false alarms. It is therefore also necessary to evaluate the fault detection algorithm for false alarm rate. The step of residual evaluation is therefore necessary for detection of fault.

7.3.1.2 Experimental Analysis

For experimental evaluation the crankshaft speed data was acquired and de-multiplexed into four data streams one corresponding to each cylinder. In actual data the de-multiplexing could have been performed by two possible methods.

- Identify the missing tooth/ double tooth and use it as reference for the next coming data
- Use igniter or injector signal along with speed signal. The signal can be observed after every six pulses in picture shown in appendix A.

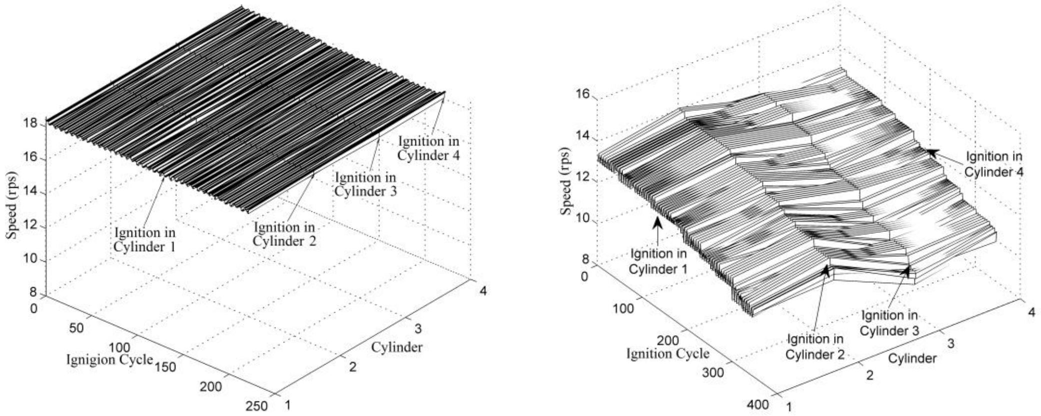


Figure 7.5 : Experimental Results: The surface representing cylinder 3 misfiring (left) no misfire (right) cylinder 3 misfiring

During experimental verification all the signals were captured but de-multiplexing was carried out using first approach only. In experiments of misfire, one of the igniter signals was blocked by pulling out its cable connector. The event is shown in picture in appendix A. This introduced a strong misfire condition in the system. As the gear has 13 teeth, only 6 data points of each sub-system could be observed so resolution of plot is one sixth fraction of power stroke.

To provide some physical insight of proposed residual, the responses of all subsystems are plotted as 3D surface shown in Figure 7.5. The 3D plot has crankshaft speed along z-axis, Cylinder number along y-axis and power stroke is plotted along x-axis. The plotted points lie on two edges parallel to axis of ignition cycle and on two slightly visible lines in between. A surface is created by joining the corresponding points on axis of cylinder number (y-axis). The smooth plot on right side of Figure 7.5 correspond to the experiment when fault was not introduced in system i.e. igniter was not removed and the non-smooth plot on left side represent experiment when fault is intentionally introduced in engine by removing an igniter.

7.3.1.3 Discussion on Results

The plots in Figure 7.5 provide useful information regarding fault diagnostic methodology.

- A residual vector corresponds to the change in height of plane from one cylinder to the next cylinder in ignition order and is characterized by maximum depth/ down-slope of surface during an ignition cycle.
- State of Markov chain can be identified by observing the cylinder number where maximum slope of an ignition strip is observed.
- Under no fault condition, the surface of 3D plot of hybrid model shown in Figure 7.5 is very smooth. The surface however lost its smoothness when misfire fault occurs as shown in Figure 3.

A quick result that visual smoothness of curve is an indication of no fault is fairly misleading as slight non-smoothness of surface is difficult to detect visually. The analysis of data in fact revealed a fault in system which was verified experimentally. The non-smoothness present in Figure 7.5 is small enough to become visible but could be identified by data analysis. Any apparent non-smoothness of surface always indicates a fault.

Under no fault condition the surface seems sufficiently smooth; however the edge parallel to time/cylinder indicates slight ripples. A zoomed view of Figure 7.5 is shown in Figure 7.6 where these ripples are more prominent. These ripples represent

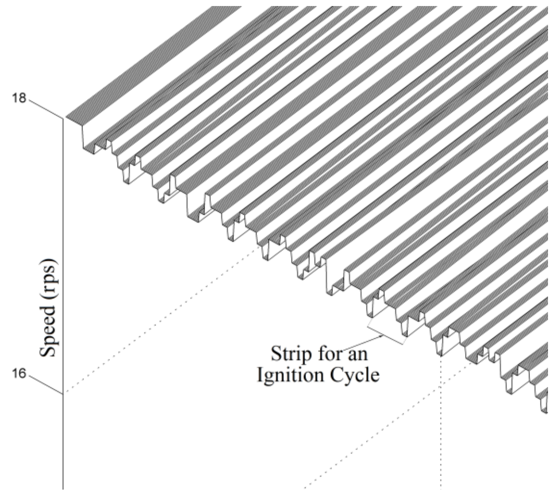


Figure 7.6 : portion of zoomed edge of surface with no misfire

power strokes of cylinders and a complete ripple strip from cylinder 1 to cylinder 4 represents a complete ignition cycle.

Both the representation of data i.e. histograms shown in Figure 7.4 and the 3D plots shown in Figure 7.5 are plotted under the major fault condition in which a spark plug was removed and no power was generated in one of the cylinder. The fault was so significant that it is visually apparent in both the figures. The occurrence of this fault was also annunciated by MIL light of vehicle. If the fault is however minor, it would not be visible in any of these figures and MIL indication of vehicle could also not capture it.

7.3.1.4 Residual Analysis using Markov Chains

The residual analysis method using Markov Chains proposed in section 6.2.4 was verified experimentally. Following are studied using experimental data:

- Validity of algorithm
- Convergence of limiting probability (by plotting probability in each iteration)
- Study of False Alarm Rate
- Comparison of method with some other existing misfire detection methods

After experimental validation of method, the possibility of application of method for the detection and identification of multiple misfire faults was studied using simulation.

Experiment 1

For experimental verification of proposed misfire detection methodology, fault was intentionally introduced in system by pulling out the connector of igniter circuit of cylinder 3. Data of only 46 ignitions cycles with fault introduced in cylinder 3 was captured. The data was analyzed according to the methodology given in section 6.2.4 by writing a program in Matlab. The results are:

$$\mathrm{F} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 35 & 1 \\ 0 & 0 & 1 & 9 \end{bmatrix} \qquad \text{(Eq 7.1)}$$

The transition probability matrix cannot be formed from the above matrix due to division of row 1 and row 2 by zero (i.e. the sum of elements of row). This restriction of F matrix is artificially removed by assigning some extra transition to all cylinders e.g. take F as a matrix with all ones rather than null matrix as initial value. When a large data set would be available for analysis, the error due to addition of these extra transitions would not only be quite negligible but also known and hence can be considered at the time of decision making.

Taking F as a matrix with all entries equal to 1 at the start, the resulting matrix after the processing of data is:

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 36 & 2 \\ 1 & 1 & 2 & 10 \end{bmatrix} \qquad \text{(Eq 7.2)}$$

$$P = \begin{bmatrix} 0.250 & 0.250 & 0.250 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.250 \\ 0.025 & 0.025 & 0.900 & 0.050 \\ 0.071 & 0.071 & 0.142 & 0.710 \end{bmatrix}$$

The eigenvalue decomposition of probability transition matrix would result in eigenvectors and a diagonal matrix as:

$$V = \begin{bmatrix} 0.7071 & -0.6787 & 0.5007 & 0.4267 \\ -0.707 & -0.6787 & 0.5007 & 0.4267 \\ 0.0000 & 0.0388 & 0.5043 & -0.3346 \\ 0.0000 & 0.2781 & 0.4944 & 0.7238 \end{bmatrix}$$

and (Eq 7.3)

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.3838 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.7305 \end{bmatrix}$$

The limiting probability is then equal to:

$$P(\infty) = [0.0645 \quad 0.0645 \quad 0.6452 \quad 0.2258] \qquad \text{(Eq 7.4)}$$

The result clearly indicates that the probability of occurrence of fault is highest in 3rd cylinder. The step of residual analysis therefore not only identified the fault with sufficient clarity but also successfully isolated the fault although a very small data sample of only 46 ignition cycles was used for analysis.

Experiment 2

The experiment was then repeated by reconnecting the igniter plug and hence removing the misfire condition. A data of 592 ignition cycles was logged on computer and analyzed. The F matrix observed as a result of analysis is:

$$\mathrm{F} = \begin{bmatrix} 294 & 10 & 6 & 3 \\ 11 & 76 & 9 & 5 \\ 5 & 12 & 105 & 4 \\ 2 & 4 & 6 & 40 \end{bmatrix} \qquad \text{(Eq 7.5)}$$

The eigenvalue decomposition of probability transition matrix is performed and diagonal matrix is shown:

$$\mathrm{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.88 & 0 & 0 \\ 0 & 0 & 0.68 & 0 \\ 0 & 0 & 0 & 0.72 \end{bmatrix} \qquad \text{(Eq 7.6)}$$

Assuming initially all the cylinders are faulty with equal probability

$$\mathrm{P}(0) = [0.25 \quad 0.25 \quad 0.25 \quad 0.25] \qquad \text{(Eq 7.7)}$$

Limiting state probability is estimated to be:

$$\mathrm{P}(\infty) = [0.5167 \quad 0.1777 \quad 0.2163 \quad 0.0893] \qquad \text{(Eq 7.8)}$$

A misfire condition in cylinder 1 is detected even when no misfire was intentionally introduced in system. The result seemed to be a false alarm condition. The results of experiment 3 however revealed that the result of experiment 2 is correct.

Experiment 3

The validity of result of experiment 2 was explored by conducting another experiment to study the air leakage from cylinders. In this experiment, all the four spark plugs were removed and pressure gauge was installed in their position. The gauges were set to retain the peak value of observed air pressure. The pistons were moved by using the starter motor. Maximum pressure created in cylinders during compression stroke was retained by pressure gauge. The observed values of cylinder pressure are given in Table 7.2. The results of this experiment indicate that slight pressure loss (misfire) was observed due to air leakage in first cylinder.

The result is promising as fault is detected when no perceptible symptoms of fault were present in engine operation. In this case the proposed algorithm has provided and ***early warning*** *of fault* that is likely to become severe in future. ECU was also not telling any fault and even the simple visual inspection of graphs in Figure 7.5 do not provide any indication of fault.

TABLE 7.2 LEAKAGE IN CYLINDERS

Cylinder	1	2	3	4
Pressure (bar)	≈ 8.75	≈ 10	≈ 10	≈ 10

The analysis of F matrice (Eq 7.2 and Eq 7.5) indicates that the algorithm has finally indicated the fault in the cylinder with largest diagonal element in F matrices. The result is again heuristic as the fault would always bias the system in a way that it will be indicated in that state.

7.3.1.5 No Misfire Condition

To study the validity of algorithm under balanced condition, the air leakage fault was removed and the experiment was again repeated.

Experiment 4

The data was acquired with all the igniter signals connected and no misfire condition present in system. The program was modified and the limiting probability vector was calculated whenever the F matrix is changed. The components of limiting probability vector were plotted with data acquisition index. The resulting plot is shown in Figure 7.7. The analysis of plot indicates the following results:

- The slightly high value of limiting probability of one engine cylinder indicates little misfire/ unbalanced condition, however the engine is operating in sufficiently balanced condition.
- The convergence of method is very fast and hence the faulty cylinder could be identified in a few ignition cycles.

The results presented in the earlier section are now being extended using analytic and simulation methods to explore the behavior or algorithm under other conditions that were not analyzed experimentally. The detail of analysis is presented in section 7.4. The experimental analysis presented in section 7.3.1.5 also clearly indicates the convergence of limiting probabilities.

Although, the results provided in section 7.3 indicate the effectiveness of proposed hybrid model and misfire detection technique. An error analysis similar to ROC analysis is however carried out to study the accuracy of proposed technique.

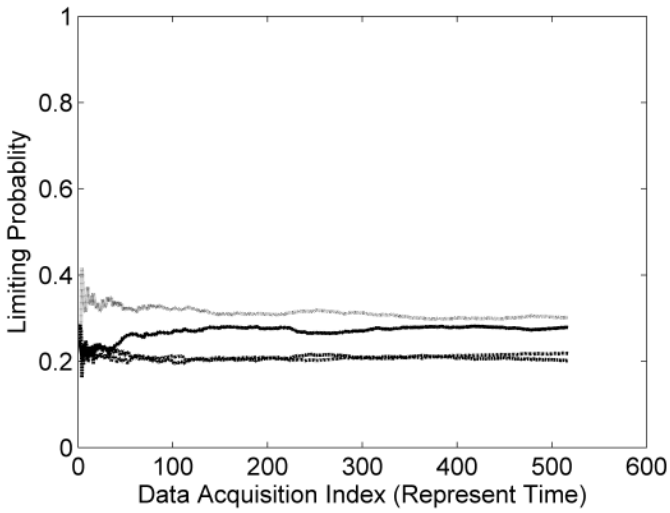


Figure 7.7 : Convergence of limiting probability with time index

7.4 Dependency of Error Rate on Number of Transitions

The error rate of the proposed algorithm is studied using an analysis similar to ROC analysis. In this analysis the predictions of algorithm for different values of "*n"* are compared with the experimental observations to identify the "True Positive" and "False Positive" predictions of algorithm. The experiment is repeated 10 times and a single point is drawn on a curve with axes identified as "True Positive Rate" and "False Positive Rate". The above plot resembles ROC curve as ROC curve is also plotted between "True Positive Rate" and "False Positive Rate". The accuracy of algorithm is estimated using the location of points on the curve. In this regard the properties of ROC analysis can be used.

7.4.1 Properties of ROC Analysis

The ROC analysis is a tool of signal detection theory to depict the tradeoff between hit rates and false alarm rate of classifiers. The properties of ROC analysis are provided in detail by Fawcett T [2006, pp: 861-874]. Swet *et al* [1988, pp: 1285–1293] mentioned that scope of ROC analysis is extended and it can be used to analyze the behavior of diagnostic systems.

When the points are plotted in ROC space, following conclusions can be drawn on the basis of appearance of points in ROC curve.

- If a point is to the north-west side of the other point in ROC space i.e. True positive rate is higher and the false positive rate is lower than it would represent a better classified point than the other point.
- The points on the left side of ROC curve close to X-axis are considered as conservative that make positive classification only with strong evidence. These points also make less false positive error. However these classifiers also exhibit less true positive rates.
- The points on the right side of ROC curve close to upper right side of curve are considered as liberal that make positive classification with weak evidence. These points classify all positive instances correctly but at the cost of very high false alarm rate.

- The diagonal line y=x represent random guessing and is known as "*Chance Line*". The random guessing will result in classification of points that repeatedly cross the Chance Line.
- A classifier is potentially optimal if and only if it lies on the convex hull of the points in ROC space.

7.4.2 Error Analysis Of Fault Diagnostic Algorithm

For error analysis, predictions were made by algorithm. The observed data set was considered as the observed instances and the algorithm was considered as the classifier to make the predictions.

A binary classification $\{pf, nf\}$ was assumed where

- *pf* represents the presence of fault in 3rd cylinder
- *nf* represents that no fault is present in 3rd cylinder

The classification was made by running the algorithm on the observed data and predicting the fault probability vector using Chapman-Kolmogorov equation:

$$p(n) = p(0)P^n \quad \text{(Eq 7.16)}$$

Probability transition matrix was estimated by analyzing data of large number of ignition cycles. To study the affectivity of algorithm, *n* was varied from 0 to some higher values to estimate the fault probability. Using the fault probability vector estimated by the algorithm predictions were made about the fault for different values of n and these predictions were compared with the experimentally observed results to identify the "*True Positive*" and "*False Positive*" events. The different values of *n* used in the above experiment correspond to the following cases.

- *n=0* means the algorithm is completely bypassed and probability vector is not updated by the processing of the available data.
- *n=2* is a better approximation of proposed algorithm where probability vector is updated by the processing of the available data.
- *n=10* is even better approximation of proposed algorithm as compare to the case of n=2.

It was observed that as n is increased to even larger values, the results of proposed algorithm become more accurate. The proposed analysis was carried out for n=0, n=2 and n=10 only. The results were then used to analyze the performance of proposed fault diagnostic algorithm.

Initially probability matrix was estimated by analysis of data set of 50 ignition cycles with fault in 3rd cylinder. The fault probability vector was estimated using n=0, n=2 and n=10 in Chapman-Kolmogorov equation. Using the fault probability vector, a new data set was generated to represent the fault estimates. The fault estimates were then compared with the original data to identify the "*True Positive*" and "*False Positive*" cases.

The experiment was then repeated with data from a maximally balance cylinder with no misfire to calculate the probability transition matrix and the fault probability was estimated using proposed algorithm. Again this fault probability was used to predict the data and compared with the experimental data to identify the "*True Positive*" and "*False Positive*" results.

Using analysis of experimental results and predictions, a point was plotted in space with "*False Positive Rate*" along X-axis and "*True Positive Rate*" along Y-axis. Ten predicted data instances that were generated corresponding to each value of n are plotted. The plot is shown in Figure 7.8. The convex hull and a chance line (Major diagonal) are also plotted on the curve for analysis. The analysis of Figure 7.8 indicates the following results:

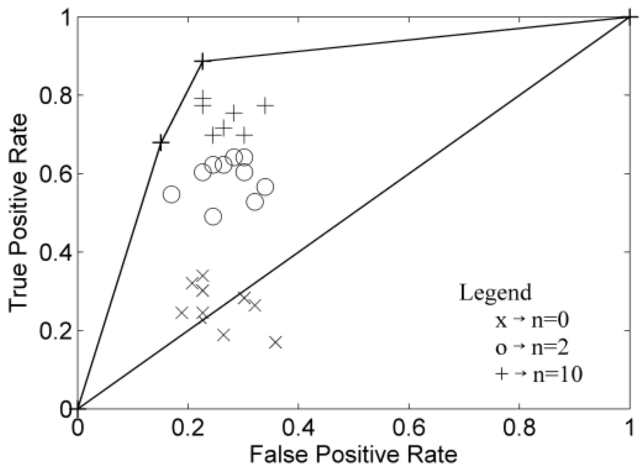


Figure 7.8 : Error Analysis of results of proposed Fault Diagnostic Algorithm

- When n=0 in Eq 7.16, i.e. when the fault diagnosis algorithm is completely bypassed, all points are close to chance line and continuously crossing it indicating a confusion state. In this case the diagnosis occurs purely on the basis of initial probabilities.
- For n=2, the cluster of points is shifted to north-west side of plot, indicating better accuracy of diagnosis algorithm.
- For n=10, the cluster of points is shifted further toward north west side and close to the convex hull, indicating even better accuracy.

An observation of results clearly indicates that the proposed algorithm is neither conservative nor liberal. Also the accuracy of fault diagnosis would improve significantly with increase in values of n. With increase in value of n, the number of false alarm would also decrease significantly. Also higher is the value of n more points would lie on the convex hull indicating more optimal behavior of algorithm. The choice of limiting probability in proposed fault diagnostic algorithm would therefore result in better detection with small false alarm rate.

7.5 Extension of Results

After experimental verification of proposed misfire detection method, some more misfire conditions were analyzed using theoretical analysis and simulation techniques. The two cases of interest are being presented:

- Response of Algorithm under Random misfire condition
- Response of Algorithm under multiple misfire condition

To study the behavior of proposed algorithm based on Markov chain, under random misfire conditions, the problem of fault detection is studied analytically. The experimental setup however needs more alterations in its electrical circuits to verify the results of random misfire mode. The experimental verification of random misfire mode is therefore not carried out.

The case of multiple misfire conditions is also explored using simulations.

7.5.1 Detection of Random Misfire Condition

Lee A. *et al* [2003, pp: 3377-3381] mentioned that random misfire mode is the most difficult fault detection problem and can be considered as a “benchmark” for misfire detection methodology. Lee A used a Kalman filtering based approach to detect the random misfire conditions. The approach was validated experimentally. A theoretical justification based on concrete mathematical ground was however not presented for the method, to justify the claim of fault detection in random misfire mode. In this section a theoretical justification based on mathematical ground is being provided to prove that the proposed algorithm can successfully detect the misfire fault even under random misfire mode.

The analysis for the detection of random misfire was carried out in the same manner as the propagation of bits in a binary communication channel which is modeled as probability of observing a bit 1 at the output of channel is α when the bit send on input to the channel is 0 and the probability of observing bit 0 at the output of channel is β when 1 is transmitted at the input.

It is already stated that under no fault condition, all the states are occurring with equal probability. However the occurrence of fault would provide an extra bias to the faulty state so that the probability of occurrence of that state would increase. The analysis also assumed that fault if present in system would be in 3rd cylinder of SI engine and the probability of fault and no fault condition at the input is same.

For random misfire analysis, the four state model of Markov chain was transformed into two state model with states G_1 and G_2. The state G_1 represents that fault is detected in 3rd cylinder and the state G_2 represents that fault is not detected in cylinder 3. When all the cylinders are healthy state G_1 occur due to false alarm. When fault is actually present in cylinder 3, the state may occur due to correct detection. Similarly when a fault is present in cylinder 3, the occurrence of state G_1 is a hit and occurrence of state G_2 is a false alarm. Since the state G_2 represents the no fault condition of cylinders 3, probability of occurrence of states would be $[0.25 \quad 0.75]$

If under no fault condition, the probability of occurrence of fault state is p then

- Probability of occurrence of state G_1 (False Alarm probability) is p
- Probability of non-occurrence of state G_1 (Hit Rate probability) is $1 - p$

Similarly, if under fault condition, the probability of occurrence of state G_2 is q where $q > p$. It is also assumed that then

- Probability of occurrence of group G_2 (False Alarm probability) is q
- Probability of occurrence of group G_1 (Hit Rate probability)is $1 - q$

The probability of occurrence of states is shown in Figure 7.9

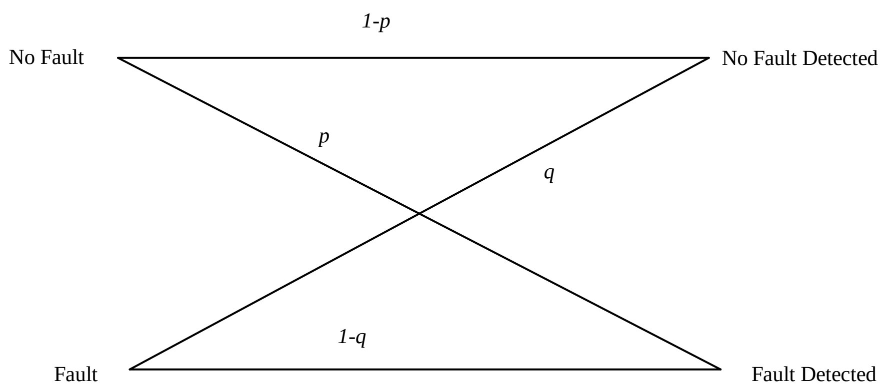


Figure 7.9 : Detection probabilities of Randomly occurring faults in SI engine

The probability transition matrix for this is given as:

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \qquad \text{(Eq 7.9)}$$

The two Eigenvalues of matrix *P* are 1 and 1-*p*-*q* and the eigenvectors are:

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad and \quad v_2 = \begin{bmatrix} -p \\ q \end{bmatrix} \qquad \text{(Eq 7.10)}$$

As per definition of Papoulis A. [2006, pp: 699-708], defining two vectors *U* and *V*, where *U* is a matrix of Eigenvector and *V* is the inverse of matrix *U*. The matrices *U* and *V* can then be used to estimate the n-step transition matrix. For the problem in hand, the matrices *U* and *V* are defined as:

$$U = \begin{bmatrix} 1 & -p \\ 1 & q \end{bmatrix} \quad and \quad V = \frac{1}{p+q}\begin{bmatrix} q & p \\ -1 & 1 \end{bmatrix} \qquad \text{(Eq 7.11)}$$

The n-step transition matrix is then found as:

$$P^n = \frac{1}{q+p}\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}q & p\end{bmatrix} + \frac{(1-p-q)^n}{q+p}\begin{bmatrix}-p\\q\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix}$$

(Eq 7.12)

$$P^n = \frac{1}{q+p}\begin{bmatrix}q & p\\q & p\end{bmatrix} + \frac{(1-p-q)^n}{q+p}\begin{bmatrix}p & -p\\-q & q\end{bmatrix}$$

Both p and q represent the events of False Alarm. Error analysis of predictions of the proposed algorithm (Figure 7.8) indicates that largest False Positive rate is 0.4 (when *n* is chosen as 10 instead of infinite) hence both *p* and *q* can be taken as 0.4.

For estimating the limiting probabilities, $n \to \infty$. As $|1-q+p| < 1$, hence for the estimation of limiting probability, the second term would be reduced to zero and Eq 7.12 would reduced to:

$$P^n = \frac{1}{q+p}\begin{bmatrix}q & p\\q & p\end{bmatrix}$$

(Eq 7.13)

The proposed algorithm would always declare fault in one of the cylinder in every ignition cycle, so the probability of declaration of fault in third cylinder would be 0.25 and in declaration of fault in any other cylinders would be 0.75. Assuming the initial probabilities of states as $[0.25 \quad 0.75]$, the limiting probabilities would become:

$$p(\infty) = \frac{1}{0.8}\begin{bmatrix}0.25 & 0.75\end{bmatrix}\begin{bmatrix}0.4 & 0.4\\0.4 & 0.4\end{bmatrix}$$

(Eq 7.14)

$$p(\infty) = \frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}$$

i.e. the probability that fault is detected in 3rd cylinder (Faulty cylinder) is 0.5 and probability that fault is not detected in 3rd cylinder is also 0.5. Assuming that when fault is detected in any other cylinder then it can belong to any cylinder with equal probability, than the probability of detection of fault in different cylinders would be:

$$p(\infty) = [0.166 \quad 0.166 \quad 0.5 \quad 0.166] \tag{Eq 7.15}$$

The limiting probability indicates that the algorithm identifies the probability of fault in cylinder 3 to be 0.5 and that of in cylinders 1, 2 and 4 is 0.166. The fault is therefore identified correctly even if random misfire fault is introduced in engine.

It is also interesting to analyze the case when the probability of occurrence of state G_1 becomes 1 under fault condition and 0.4 under no fault condition on the basis of ROC analysis. Here false positive case is emphasized for correct detection but false alarms are allowed in no fault condition. In this case $q = 1$ and $p = 0.4$ and the value of limiting probability becomes:

$$p(\infty) = [0.714 \quad 0.285] \tag{Eq 7.15}$$

The above analysis clearly indicates the following:

- The limiting probability would converge to a fixed value depending upon the intensity of fault condition.
- Random misfire conditions can be detected successfully even when the intensity of misfire condition is not very high.

To simulate the random misfire a condition, the model was initially tuned by adjusting the gains of four sub-systems so that the value of limiting probability remain as close to [0.25 0.25 0.25 0.25] as possible. A random number with a mean value 0 and variance 1 was generated but latched only once in each ignition cycle. The number was compared with 0 (Threshold) and if the number is found greater than zero, the gain of faulty cylinder is adjusted to 1 and otherwise is set to 0. In this way random misfire events were simulated with misfire probability of 0.5. The probability of misfire events were changed by changing the threshold value from zero to some other values e.g. choosing threshold to be less than -1, implies misfire event will almost never occur and choosing threshold to be greater than 1 implies that misfire event would occur almost certainly. The condition of random misfire was simulated in Matlab/ Simulink using proposed hybrid model. The simulation block diagram is shown in Fig. 7.10.

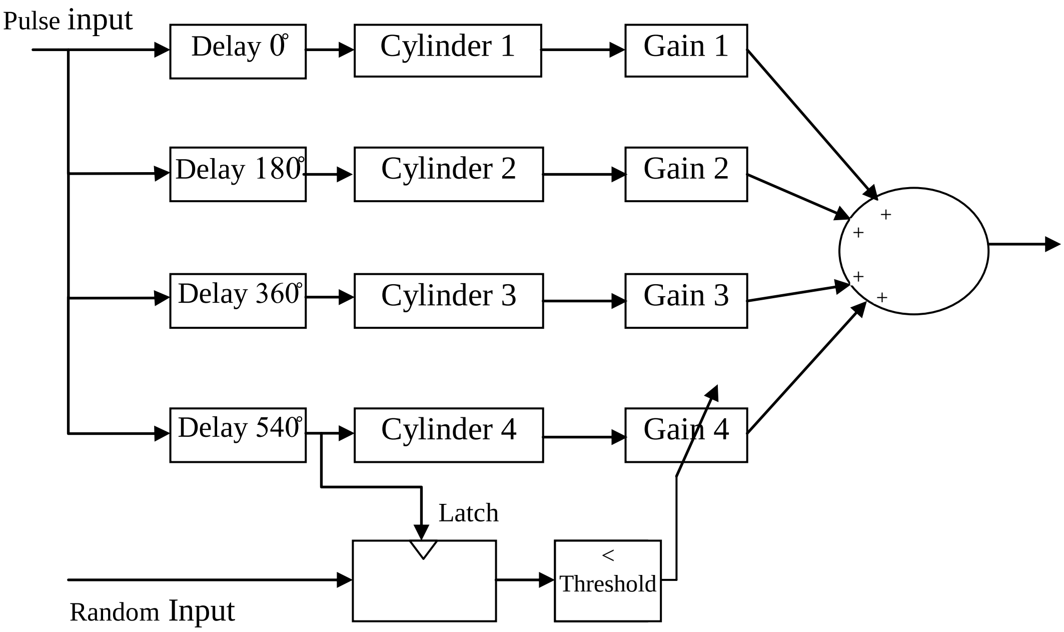


Figure 7.10 : Block Diagram of Experiment of Random Misfire

Two cases of predictions of algorithm, one with 50% random misfire in cylinder 1 and other with 50% random misfire in cylinder 2 are provided in Table 7.3.

TABLE 7.3 SIMULATION RESULTS OF RANDOM MISFIRE

Random Fault	Limiting Probability
No Fault	[0.2695 0.2192 0.2527 0.2586]
Cylinder 1	[0.6032 0.1293 0.1244 0.1431]
Cylinder 2	[0.1125 0.6190 0.1254 0.1431]

7.5.2 Detection of Multiple Misfire Events

The study of multiple misfire events was carried out using simulation methods. In this regard data was generated using hybrid model. Random speed fluctuations as observed in experimental data were introduced in the output of the system by adding small noise in data. The conditions of both no-misfire and different misfire events were simulated in model. The data was saved in a file for further analysis by the proposed misfire detection algorithm. The results indicating the limiting probability vector are provided in Table 7.5.

TABLE 7.4 SIMULATION RESULTS OF MULTIPLE MISFIRE DETECTION

Fault	**Limiting Probability**
No Fault	[0.2556 0.2181 0.2441 0.2823]
Cylinder 1	[0.9647 0.0118 0.0118 0.0118]
Cylinder 2	[0.0099 0.9703 0.0099 0.0099]
Double Fault Cylinder 1 +3	[0.5088 0.0118 0.4676 0.0118]

The results indicate that under no misfire condition the probability of fault remain almost same for all cylinders but under misfire conditions the probability of fault for misfiring cylinder, become larger. Under double-misfire conditions, the probability increases for both the misfiring cylinders.

The combined analytic, experimental and simulation results indicate that the proposed algorithms can detect incipient faults as well as intermittent faults. The method can be extended for the detection of multiple misfire. The method can also be used to generate early warnings of faults.

7.5.3 Early Warning

Early warnings not only provide prior information of faults likely to occur in future but can also be used to estimate the *Remaining Useful Life* (RUL) of a product. For estimation of RUL the intensity of detected incipient fault has to be established. A good estimate of RUL can improve the *Reliability* of product.

The results of Experiment 2, described in section 7.3.1.4 are quite promising as no perceptible evidences of misfire are evident in data plotted in Figure 7.5. Engine ECU was also not providing any indication of misfire condition, but the algorithm has successfully detected the misfire condition. The results of cylinder leakage indicate almost 10% air leakage from cylinders which is a slight misfire condition. This indicates that the algorithm can successfully detect the slight misfire faults also. With

The experimental setup did not provide facility to introduce controlled amount of fault. Simulation technique is therefore used to analyze the behavior of algorithm under different fault intensities.

The proposed hybrid model was used to study the value of limiting probability of Markov chains under varying fault conditions. In this regard fault was introduced in one of the cylinder and severity of fault is gradually increased. The model generated for study of random misfire condition was used in this case because the severity of fault can be increased simply by adjusting the threshold value in Fig. 7.10. A threshold value of 0 means 50% misfire events are being occurring. A threshold value of 0.2 means almost 60% probability of misfire etc. The data generated using model was then analyzed to identify the values of limiting probability under different misfire conditions.

TABLE 7.5 SIMULATION RESULTS OF GRADUAL RISE OF FAULT

Random Fault	**Limiting Probability**
100% Misfire	[0.9872 0.0039 0.0049 0.0039]
90% Misfire	[0.8264 0.0602 0.0483 0.0652]
80% Misfire	[0.7730 0.0759 0.0651 0. 0859]
70% Misfire	[0.7226 0.0918 0.0829 0. 1027]
60% Misfire	[0.6693 0.1086 0.1027 0. 1194]
50% Misfire	[0.6032 0.1293 0.1244 0. 1431]
40% Misfire	[0.5410 0.1461 0.1510 0. 1619]
30% Misfire	[0.4738 0.1678 0.1737 0. 1846]
20% Misfire	[0.4284 0.1856 0.1846 0. 2014]
10% Misfire	[0.3860 0.1984 0.1974 0. 2182]

In simulations misfire condition was introduced in first cylinder and the intensity of fault was increased in the steps of 10% by changing the value of threshold from -1 to +1. The results of data analysis under different misfire conditions are provided in Table 7.4.

The simulation results provided in Table 7.4 indicate that the limiting probability of misfiring cylinder is increasing almost linearly with fault. The wear out region would not be sharp and a safe value of threshold can be defined on the basis of simulation results for generating early warning of fault. The warning would be generated well before the occurrence of complete failure.

7.5.4 Misfire Detection as Reliability Problem

The detection and identification of incipient misfire condition lead us to the area of prognosis. A brief terminology used in prognosis include "Remaining Useful Life" RUL and Reliability. These areas represent the possible future extension of work presented in this thesis. To provide some idea of future extension, mathematical formulation of reliability is applied to Misfire fault detection problem.

A "*Reliability Function*" could be generated to provide current intensity of fault. By identifying the location of point on bath tub curve, an educated guess can be made whether the operation is in useful life region or in wear out region. The basic terminology and definitions about reliability function can be found from Pradhan D. K. [1995, pp: 55-57].

The proposed fault diagnostic method works on the basis that one of the cylinders is assumed faulty in each ignition cycle. If a particular cylinder was declared faulty $N_f(t)$ times in N ignition cycles, then reliability of the cylinder at time t would be given as:

$$R(t) = 1 - \frac{N_f(t)}{N}$$

Reliability function is defined as a differentiation of *R(t)* with respect to time.

$$\frac{dR(t)}{dt} = -\frac{1}{N}\frac{dN_f(t)}{dt}$$

$$\frac{dN_f(t)}{dt} = -N.\frac{dR}{dt}$$

That represents the instantaneous rate at which the cylinder is declared as faulty. If

$$N = N_0 + N_f$$

i.e. if cylinder was declared faulty N_f times then it was declared healthy N_0 time. The "*Hazard Function*" or the "*Failure Rate Function*" is then defined as:

$$z(t) = \frac{1}{N_0(t)} \frac{dN_f(t)}{dt}$$

If a test of N ignition cycles is conducted, then frequent visit of any state would not only increase the $\frac{dN_f(t)}{dt}$ but also decrease the value of $N_0(t)$, resulting in increase of hazard function. The typical curve of hazard function is called *"bathtub curve"* with three distinct regions. In the first region the value of hazard function is high and decreases rapidly. This region is called the infant mortality region and is considered to come in the time before start of the useful life of component. In second region the curve is sufficiently flat and is considered as the useful life period of component. In the third region the curve rises rapidly and is known as the wear out region. A threshold value can not only be defined to generate an early warning before the failure but also can be used to identify the remaining useful life.

7.6 Comparison Of Method

A range of different evaluation criteria were used for the performance evaluation of fault diagnostic algorithms. A framework to compare and evaluate the diagnostic algorithms was created jointly by "*NASA Ames Research Center*" and "*Palo Alto Research Center - PARC*" for a competition called *DXC'09* [Kurtoglu T *et al* , 2009]. Following evaluation criteria were adopted for the comparison of different diagnostic algorithms.

- False Positive Rate : indicating spurious faults rate
- False Negative Rate : indicating missed faults rate

- Detection Accuracy : indicate correctness of detection
- Fault Detection Time : indicate time for detecting a fault
- Fault Isolation Time : Time for last persistent diagnosis
- Diagnostic Utility : indicate cost related to component replacement due to incorrect diagnosis
- CPU load : indicate CPU time spent for diagnosis
- Memory load : indicate the memory allocation

Error analysis of proposed algorithm is already presented in section 7.4. Result of experiment 4 and Figure 7.7 indicates very small fault detection time. For most of the fault diagnostic algorithms found in literature such analysis for False Positive Rate or Detection Accuracy etc. are not provided for comparison. By the analysis of the algorithms provided in literature it is however possible to identify the CPU load and Memory load of algorithms. The proposed fault detection algorithm would therefore be compared to some other misfire detection algorithms found in literature on the basis of memory requirements (Memory load) and number of computations (CPU load) in a complete ignition cycle.

The literature survey on detection of misfire fault in SI engine provided in chapter 3, indicates a wide range of method adopted for the detection and isolation problem. The proposed algorithm would be compared with methods based on cross-correlation of the observed signal with the signals of known faults. Correlation technique was used in a number of different proposals like Sood A. K. *et al* [1985, pp. 294-300] and Rizzoni, G. *et al* [1988, pp: 237-244] etc.

Consider a vector with N samples in complete ignition cycle. The cross-correlation coefficient of data vector with known fault vector is given by:

$$S_{ij} = \frac{\sum_{n=1}^{N}[(x_i(n) - \hat{x}_i)(x_j(n) - \hat{x}_j)]}{\sigma_i \sigma_j} \qquad \text{(Eq 7.17)}$$

7.6.1 Memory Load

The method based on cross correlation need N memory locations for data and N memory location for each sample of known faulty signal. If N = 24 then 48 memory locations would be needed to detect a *single* fault. The proposed method need only 25 memory locations: a system counter, four counters of subsystems, four igniter signals and 16 elements of state transition matrix. The method can detect all single fault cases.

7.6.2 CPU Load

The method based on correlation will first calculate velocity vector. The average and standard deviations of velocity vectors are computed. The correlation coefficient would then be found by N multiplications, N additions and a division for each fault. The proposed method needs only N+8 comparisons and 12 additions to identify all faults.

The other methods based on model based fault detection and wavelet based techniques need floating point calculations and is computationally more expensive. The implementation philosophy and flow chart of proposed algorithm indicates the simplicity of proposed fault diagnosis algorithm without floating point calculations.

7.7 Summary

This chapter presented the results and discussed them in details. The results section is divided into two basic areas i.e. the validation of model and validation of fault diagnostic algorithm. The chapter is started with a brief description of experimental setup, its development and its weaknesses. The proposed hybrid model was programmed using Matlab. The values of different coefficients were selected on the basis of engineering judgment, knowledge of experts and data observed using internet. The simulation results were then compared with the experimental data and the values of different model parameters were tuned to match the simulation results as closely to the experimental results as possible under no fault condition. When the simulation results matched the experimental results, misfire fault was introduced first in experimental setup and the faulty data is recorded. Misfire fault was then introduced in the simulation and the response of model under faulty condition was observed. The results of experimental and simulation under faulty conditions were then compared to validate the proposed model.

The experimental results under no-fault conditions and faulty conditions were then used in the proposed fault detection algorithm to identify the fault. The results of simulation and experiments were plotted in a number of different ways to identify the different events and steps of fault diagnostic algorithm. In this regard a histogram plot of crankshaft speed variations indicated the residual. A 3D plot of engine speed, ignition cycle and cylinder number provided the visual indication of ignition cycle, residual etc. The experimental results under healthy engine condition and in the presence of fault were then finally processed using the proposed fault diagnostic algorithm. It was concluded that the fault is successfully identified from the data of experiment when fault was present in engine. It was however observed that the data of healthy engine also indicate the fault. The setup was then explored for any possible fault in it and a new experiment was conducted in which the spark plugs were removed and a pressure indicator was placed in their position. When the cylinder pistons were moved by rotating the starting (self) motor, pressure was established in different cylinders. It was observed that air leak fault was observed in one of the engine cylinder indicating slight misfire condition.

The experiment was later repeated when engine was operating under healthy conditions and it was observed that no fault was indicated by the algorithm. To study the response of algorithm for data analysis, simulations were performed without predictions, two step prediction and ten step prediction using Markov chains (n=0, n=2 and n=10 respectively in Eq 6.16). The results were then used to predict the faults in engine system and the engine data was compared with the predicted data to identify the false positive and true positive rates. An ROC curve was plotted for all these observations and it was observed that the results corresponding to n=0, i.e. when the algorithm was totally bypassed were totally random and lie close to the confusion line on the plot. As the data was predicted using Markov assumption, the results were improved. In this regard the results corresponding to n=10 were found to be the best results that lie close to the convex hull of the ROC curve. ROC plot clearly indicated that the results corresponding to the limiting probabilities would be close to the experimental results and thus the fault can be identified using the mentioned algorithm.

The validity of algorithm was then explored for the intermittent misfire fault condition. The basic problem in this regard was the validation of results as the experimental setup provided no method to introduce random faults in it. The problem of validity of algorithm for intermittent misfire fault was therefore explored analytically and it was proved that the algorithm is capable to detect intermittent misfire faults also. The analytic results were then validated using simulations by representing SI engine by Hybrid model and introducing random misfire fault in it. The simulation results also indicated that the algorithm can be used for the detection of intermittent misfire fault condition, incipient fault conditions and misfire fault in multiple cylinders.

The problem of misfire detection is also posed as a reliability problem. This work can be extended in future to develop a method for prognosis by finding the position on bathtub curve.

The chapter is finally concluded by comparing the proposed misfire fault diagnostic algorithm with some other algorithms proposed in literature for detecting misfire fault on the basis of amount of memory needed and number of computations required by different algorithms.

Chapter 8

CONCLUSIONS AND FUTURE WORK

This thesis was aimed to contribute to Misfire Fault Detection in Spark Ignition Engine. This chapter provides a brief overview of basic motivating factor behind the work and its methodology. The main contributions of work are presented in section 8.1. A brief summary of future works is described in section 8.3 and the social contribution of research work and development work carried out during the study and execution of this research is described in section 8.4.

8.1 Main Contributions

The main contribution of the thesis can be distributed in two major areas:

- Hybrid Modeling of SI engine
- Early detection of misfire fault in SI engine

8.1.1 Hybrid Modeling of SI Engine

The development of Hybrid model was motivated on the basis that SI engine contain both continuous and discrete states. Also for application of Markov Chains it was desired that output of model be defined so that it could easily be transformed to states for analysis. Since geometry of SI engine indicates limited number of cylinders, special emphasis was given to the geometry of SI engine in developing the model. In deriving the engine model for a four cylinder SI engine, model of each of the four cylinders was derived and linked with the discrete event representing the spark signal that make a cylinder active to deliver power. The mathematical abstraction was established as a model combining the discrete state representing the production of spark and continuous state as the evolution of crankshaft speed. In context with the proposed model following studies and validations were established.

- The simulated model response was validated by performing experiments on SI engine of production vehicle.

- The properties of model were explored to establish some experimental observations and facilitate the statistical analysis of output variable of proposed hybrid model.
    - One basic result was that under steady state condition, all the sub-systems (cylinders) of SI engine are decoupled. This result was used to make the basic assumption of independence of events in statistical analysis of engine response.

After development of novel hybrid model for SI engine, statistical properties of crankshaft speed variations were studied and the following results were established:

- It was established that the variable representing air sucked in engine cylinder is Gaussian using "Central Limit Theorem"
- It was established that random process representing speed variations observed on the crankshaft can be considered as Gaussian and Markov process using properties of proposed hybrid model.

8.1.2 Early detection of misfire fault in SI engine

Using properties of hybrid model a set of four states were formed using output variable of hybrid model. Each state represented faulty response of one of the engine cylinder i.e. the presence of system in state was a representative of the presence of fault in cylinder corresponding to that state. A state transition matrix representing the jumps from one state to the next state was defined. It was assumed that initially the probability of fault in all engine cylinders is equal. An algorithm based on the estimation of limiting probability was proposed to estimate the fault.

The basic strength of algorithm was observed when the MIL light of engine was indicating no fault condition and apparent behavior of engine was also indicating a healthy behavior but the algorithm detected a fault and later the experimentation indicated a minor air leakage fault in one of the engine cylinder. This indicated that the proposed algorithm could be used to generate the early warning in a system. However intensity of fault could not be defined due to unavailability of appropriate experimental setup.

The algorithm was validated by conducting experiments on the engine of a production vehicle. To analyze the performance of algorithm following studies were carried out:

- Algorithm was tested for the detection of multiple misfire events using simulations
- The benchmark problem of detection of multiple misfire events was studied using theoretical calculations
- ROC analysis of estimation was carried out to study the "True positive" and "False positive" rate by estimating faults using proposed algorithm and experimental data.
- It was proved analytically that the proposed misfire detection algorithm is capable to detect the random misfire condition or intermittent misfire fault also. In this regard data under random misfire conditions was generated using proposed hybrid model and response of algorithm to detect the random misfire condition was tested.
- The algorithm was finally compared with correlation based misfire fault detection algorithm on the basis of number of computation involved and memory requirements of algorithm.

8.3 Future Work

A little exploration of work presented in this thesis clearly indicates the following new directions of research that start from the work presented in this thesis:

- The proposed hybrid model of SI engine is valid for steady state operation of SI engine only. The nature of SI engine is hybrid with both continuous and discrete states. It is therefore possible to develop a generalized hybrid model for SI engine valid for both transient and steady state engine operation

- The fault diagnostic methodology proposed on the basis of Markov chain is studied at very high data rates so that exact waveforms of crankshaft signal can be formed and peaks could be detected. The analysis of data on the basis of hybrid model at the ECU rate can also detect the misfire problem when the tools of stochastic analysis would be applied appropriately. The study and development of this method is another extension of the presented work.

- The alteration of setup to control the fault intensity and use the setup to identify the threshold conditions for "Early Warning Generation" and "Prognosis Applications"

- Development of model based fault detection techniques using proposed hybrid model and detecting fault using techniques like sliding mode observer, Kalman filters or Linear Matrix Inequalities (LMI).
- The potential of "Artificial Neural Network" and "Fuzzy Logics" can be used to develop new methods for the detection of misfire fault in SI engine.

8.4 Social Contribution of Research

The setup was also established during the course of research by writing an R&D proposal from ICT R&D Fund. The contract document indicates details of work committed in the project [Mohammad Ali Jinnah University, Project Funding Agreement, 2009]. The setup was obtained at the cost of some extra work of "Development of a Low Cost Fault Diagnostic Toolkit for Honda vehicle" for automotive mechanics. The social impact of training of some mechanics using those toolkits can therefore be attributed as an indirect social contribution of this project.

The literature survey for the development was carried out in parallel with the literature survey of problem proposed in this thesis. Initially the diagnostic toolkit was developed using ELM-327 IC that provides the protocol conversion from ECU protocols to RS-232. The work was however continued to interface microcontroller with vehicle ECU using ISO-9141-2 protocols and its variants. Finally the hardware and software was developed to interface microcontroller with ECU using CAN protocol. The information about the new trends of vehicle diagnostics, sensors and actuator present in vehicles and different protocols used in vehicles was provided to the mechanics working on the device. This resulted in some knowledge uplifting of those mechanics.

The availability of technology also created the opportunities of locally fabricating the low cost vehicle fault diagnostic kits for the local mechanics if sufficient finances could be made available. This would further increase the indirect beneficiaries of this research.

Albertoni L, Balluchi A, Casavola A, Gambelli C, Mosca E, Vincentelli A. L. S, "Idle Speed Control of GDI Engines using Robust Multirate hybrid Command Governors", Proceedings of IEEE Conference on Control Applications, Vol 1, 2003, pp: 140-145.

Angeli C., "On-Line Fault Detection Techniques for Technical Systems: A Survey", International Journal of Computer Science & Applications, Vol. I, No. 1, pp. 12 – 30, 2004

Aono T, Fukuchi E, "Misfire Detection Method robust against Crankshaft Vibration and Acceleration", IEEE Conference on Control Applications CCA, Toronto 2005, pp: 1218-1221.

Arogeti S. A, Wang D, Low C. B, "Mode Identification of Hybrid Systems in the Presence of Fault", "IEEE Transactions on Industrial Electronics, Vol 57, No. 4, 2010, pp: 1452-1467

Ball J. K, Bowe M. J, Stone C. R, McFadden P. D, "Torque Estimation and Misfire Detection using Block Angular Acceleration", SAE 2000-01-0560.

Balluchi A, Benvenuti L, Benedetto M. D. *etal*, "Automotive Engine Control and Hybrid Systems: Challenges and Opportunities", Proceedings of the IEEE, Vol 88, No. 7, pp. 888-912, July 2000

Basseville, M. (1988), "Detecting Changes in signals and systems - a survey," Automatica, Vol.24, No. 3, pp.309-326

Beard, Failure Accommodation in Linear Systems through Self-Reorganization. PhD dissertation, MIT, 1971, pp: 54-56

Biswas G., Marie-Odile Cardier, Jan Lunze, Louise Trave-Massuyes, and Marcel Staroswiecki, "Diagnosis of Complex Systems: Bridging the Methodologies of FDI and DX Communities", IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol, 34, No 5, October 2004 pp 2159-2162

Bohn C., Magnor O., Schultalbers M. "State Observer based Analysis of Crankshaft Speed Measurements with Application to Misfire Detection", International Conference on Control and Automation (ICCA 2005), Budapest, Hungary, 2005, pp: 239-244

Bohn C, Magnor O, Schultalbers, "State Observers for Periodic Signals: A Case Study in Misfire Detection", "Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, Volume 221, number 5, 2007, ISSN-0954-4070, pp: 641-649.

Branicky M. S, "Multiple Lyapunov Functions and Other Analysis Tools for Switched and Hybrid Systems", IEEE Transactions on Automatic Control, Vol 43, No. 4, April 1998, pp: 475-482.

Brotherton T, Jahns G, Jacobs J, Wroblewski D. "Prognosis of Faults in Gas Turbine Engines", Aerospace Conference Proceedings, 2000 IEEE, Vol 6., pp: 163-171,

Butt. Q. R and A. I. Bhatti, "Estimation of Gasoline Engine Parameters using Higher Order Sliding Mode", IEEE Transaction on Industrial Electronics Vol 55 Issue 11 Nov 2008., Page: 3891-3898

Butt Q. R, A. I. Bhatti, M. Iqbal, M. A. Rizvi, R. Mufti and I. H. Kazmi, “Estimation of Automotive Engine Parameters: Part I: Discharge Coefficient of Throttle Body”, presented in IBCAST 2009, held in January 2009, Islamabad Pakistan

California Environmental Protection Agency, Air Resources Board (CARB), “Staff Report: Initial Statement of Reasons for Proposed Rulemaking; Technical Status and Revisions to Malfunction and Diagnostic System Requirements for Heavy-Duty Engines (HD OBD) and Passenger Cars, Light-Duty,Trucks, and MediumDuty Vehicles and Engines (OBD II), 2009, pp: 117.

Chen, J. and Patton, R. J., (1999), "Robust Model-based Fault Diagnosis for Dynamic Systems", Kluwer Academic Publishers, 1999

Cheol-Kwan Yang and Duk-Sun Shim, “Double Faults Isolation Based on the Reduced-Order Parity Vectors in Redundant Sensor Configuration”, International Journal of Control, Automation, and Systems, vol. 5, no. 2, pp. 155-160, April 2007

Chow, E.Y. and Willsky, A., (1984), “ Analytical Redundancy and the Design of Robust Failure Detection Systems, ” IEEE Trans. Automat. Contr., Vol. 29, No.7, pp. 603-614

Chung Y, Choi S, Bae C, Yoon K, “A New Concept of Misfire Detection using a Wide-Range Oxygen Sensor in a Spark Ignition Engine”, Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering, Volume 213, Number 6, 1999, pp: 585-594(10).

Clark, R.N., (1989), “ State Estimation Schemes for Instrument Fault Detection, ” Fault Diagnosis in Dynamic Systems: Theory and Application, edited by Patton R.J., Frank P.M. and Clark R.N., Englewood Cliffs, NJ: Prentice Hall

Claudia M., William Maul, Michael Binder,” Rocket Engine Diagnostics Using Qualitative Modeling Technique”, Report No:NASA/TM-105714, 1992

Cook J. A, Kolmanovski I. V, McNamara D, Nelson E. C, Prasad K. V, “Control, Computing and Communication: Technologies for the Twenty First Century Model T”, Proceedings of the IEEE, Vol 95, No. 2, Feb 2007, Digital Object Identifier: 10.1109/JPROC.2007.888384, pp: 334-355.

Crossman J. A, Guo H., Murphy Y. L, Cardilo J., “Automotive Signal Fault Diagnosis-Part I: Signal Fault Analysis, Signal Segmentation, Feature Extraction and Quasi Optimal Feature Selection”, IEEE Transactions on Vehicular Technology, Vol 52, No. 4, July 2003, pp:1063-1075

Deligiannis V, Manesis,S : “Modeling Internal Combustion Engines Using a Hyper-Class of Hybrid Automata: A Case Study”, Conference on Computer Aided Control Systems Design, Munich, Germany, Proceedings of the 2006 IEEE, pp. 2991-2996

Devasenapati S. B, Ramachandran K. I, Sugumaran V, “Misfire Detection in Spark Ignition Engine using Support Vector Machines”, International Journal of Computer Applications, Vol 5, Number 6, August 2010, pp: 25-29.

Dexter, A.L. and Benouarets, M., (1997), “ Model-based Fault Diagnosis Using Fuzzy Matching, ” IEEE Trans. on Sys. Man. and Cyber. Part A: Sys. & Humans, Vol.27, No.5, pp.673-682

Ding S. X., “Model-based Fault Diagnosis Techniques Design Schemes, Algorithms, and Tools”, Springer-Verlag Berlin Heidelberg, 2008, ISBN 978-3-540-76303-1.

Ding S. X, Yang G, Zhang P, Ding E. L, Jeinsch T, Weinhold N, Schultalbers M, “Feedback Control Structures, Embedded Residual Signals, and Feedback Control Schemes With an Integrated Residual Access”, IEEE Transactions on Control Systems Technology, 2009. Digital Object Identifier 10.1109/TCST.2009.2018451, pp:1-16

Ehsan S. T., Khorasani K., “Fault Diagnosis of Nonlinear System using a Hybrid Approach”, Springer science+business Media, 2009, , ISBN 978-0-387-92906-4

Emani-Naeini A., Akhter M.M. and Rock S.M., (1988), “ Effect of Model Uncertainty on Failure Detection- The Threshold Selector, ” IEEE Trans. Automat. Contr., Vol. 33, pp. 1106-1115

Engell S, Frehse G, Schneider E, “Modeling Analysis and Design of Hybrid Systems”, Springer Verlag Berlin, Heidelberg, NewYork 2002, ISBN: 3-540-43812-2

Fahmida. M. Chowdhury, B. Jiang, C. M. Belcastro, “Reduction of False Alarms in Fault Detection Problems” International Journal of Innovative Computing, Information and Control Vol. 2, Number 3, June 2006 pp. 481-490

Fawcett T., “An introduction to ROC analysis”, Pattern Recognition Letters 27, 2006, pp. 861-874

Feldkamp. L. A, Feldkamp T. M, Prokhorov D. V. “Adaptive Classification”, Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC, IEEE 2000, pp: 52-57

Franco J, Franchek M. A, Grigoriadis K, “Real-time brake torque estimation for internal combustion engines”, Mechanical Systems and Signal processing, 22 (2008), pp: 338-361

George Vachtsevanos, Frank Lewis, “Intelligent Fault Diagnosis and Prognosis for Engineering Systems”, John Wiley and Sons, 2006. ISBN-13: 978-0-0471-72999-0.

Gertler, Janos J., (1988), "Survey of Model-based Failure Detection and Isolation in Complex Plants", IEEE Control Systems Magazine, Vol. 8, Issue:6, pp.3-11

Gertler J, “All linear methods are equal--and extendible to (some) nonlinearities”, International Journal of Robust and Nonlinear Control, 2002.

Gertler J, Jin Cao, “Design of optimal structured residuals from partial principal component models for fault diagnosis in linear systems”, Journal of Process Control, 2004

Gertler J, “Survey of model-based failure detection and isolation in complex plants,” IEEE Control Systems Magazine, vol. 3, pp. 3—11, 1988.

Gertler J.,Jin Cao, “PCA-Based Fault Diagnosis in the Presence of Control and Dynamics”, Published online 00 Month 2004 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/aic.00000

Giorgetti N, Ripaccioli G, Bemporad A, “Hybrid Model Predictive Control of Direct Injection Stratified Charge Engines”, IEEE/ASME Transactions on Mechatronics, Vol 2, No. 5, pp. 499-506, Oct 2006.

Guzzella L, Onder C. H, “Introduction to Modeling and Control of Internal Combustion Engine Systems” Springer Verlag Berlin Heidelberg 2004, ISBN 3-540-22274-x

Hendrick E. , Sorenson S. C., “Mean Value Modeling of Spark Ignition Engines” SAE Technical Paper no. 900616, 1990.

Hoffling T., R. Isermann, “Fault Detection based on Adaptive Parity Equations and Single Parameter Tracking”, Control Engineering Practice, Vol 4, No. , 1996. pp: 1361-1369.

Holzmann H., Nelles O., Halfmann C., Isermann R., “Vehicle Dynamics Simulation Based on Hybrid Modeling”, Proceedings of the 1999 IEEE/ ASME International Conference on Advanced Intelligent Mechatronics, September 19-23, 1999, Atlanta, USA. pp: 1014-1019)

Huaqun G. “Automotive Informatics and Communicative Systems: Principles in Vehicular Networks and Data Exchange”, Information Science Reference (an imprint of IGI Global), Copyright © 2009 by IGI Global ISBN 978-1-60566-338-8 (hardcover)

Iqbal M., A. I. Bhatti, S. Iqbal, Q Khan, “Robust Parameter Estimation of Nonlinear Systems using Sliding Mode Differentiator Observer” , (Accepted for publication in IEEE Transaction on Industrial Electronics, February 2010)

Isermann, “Model Based Fault Detection and Diagnosis- Status and Applications”, Annual Reviews in Control 29 (2005) 71–85

Isermann R., P. Ballé: “Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes”. Control Engineering Practice, **5**(5), 1992, pp: 709-719,

Isermann R, Müller N., “Modeling and Adaptive Control of Combustion Engines With Fast Neural Networks”, European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation on Smart Adaptive Systems, Tenerife, Spain, 2001.

Isreal K., Krishna M. “Fault Tolerant Systems”, Elsevier 2007, ISBN 18-978-0-12-088568-8.

Jacobson, C. A. and Nett, C.N., (1991), “ An Integrated Approach to Controls and Diagnostics using the Four Parameter Control, ” IEEE Contr. Syst. Mag., Vol. 11, No.6, pp. 22-29

Jean-J., E. Slotine, Weiping Li, “Applied Nonlinear Control”, Prentice Hall Inc. 1991, ISBN 0-13-040890-5

Jianhui L, Pattipati K. R, Qiao L, Chigusa S, “An Integrated Diagnostic Development Process for Automotive Engine Control Systems”, IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews, Volume 37, Number 6, November 2007, pp: 1163-1173.

Jianhui L, Madhavi N, Pattipati K. R, Qiao L, Chigsua S, “Integrated Model Based and Data Driven Diagnosis of Automotive Antilock Braking System”, IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans, 2009, pp: 1-16.

Kami´nski T., M. Wendeker, K. Urbanowicz, G. Litak, “Combustion Process in a Spark Ignition Engine: Dynamics and Noise Level Estimation”, American Institute of Physics, 2004

Karisson J, Fredriksson J. “Cylinder-by-Cylinder Engine Models Vs Mean Value Engine Models for use in Powertrain Control Applications”, Society of Automotive Engineers, SAE 1998. 99P-174, pp: 1-8

Kiencke U., L. Nielson, “Automotive Control Systems”, ISBN 3-540-23139-0, Springer Berlin Heidelberg New York, 2005

Kim D, Park J, “Application of Adaptive Control to the Fluctuation of Engine Speed at Idle”, Information Science: an International Journal, Vol 177, Issue 16, 2007, pp: 3341-3355.

Kim Y. W, Rizzoni G. Utkin V. “Automotive Engine Diagnosis and Control via Nonlinear Estimation”, IEEE control systems, 1998, pp: 84-99

Kurtoglu T, Narasimhan S, Poll S, Garcia D, Kuhn L, Kleer J, Gemund A. V, Feldman A, “First International Diagnosis Competition – DXC’09”,

Lee A., Robert N. K., Wu Z. J., “Automotive Engine Misfire Detection using Kalman Filtering” 0-7803-7954-3/03, 2003 IEEE, pp: 3377-3381.

Lee M, Yoon M, Sunwoo M. Park S, Lee K, “Development of a New Misfire Detection System Using Neural Network”, International Journal of Automotive Technology, Vol 7, Number 5, 2006, pp: 637-644.

Liberzon D. “Switching in Systems and Control”, © 2003 Birkhäuser Boston, ISBN 0-8176-4297-8

Malaczynski G. W, Van der Poel R, “Phase Diagrams of Different Modes of Misfire Calculated from the Digital Fourier Transformation of Angular Crankshaft Velocity” SAE DOI: 10.4271/2010-01-0167.

Manikandan V., Devarajan N., “Mathematical Approach towards Fault Detection and Isolation of Linear Dynamical Systems”, International Journal of Mathematics Sciences Volume 1 Number 1,2007, page:82-91

Mander E. J., Barford L. A, Biswas G, „”An Approach for Fault Detection and Isolation in Dynamic Systems From Distributed Measurements”. IEEE Transactions on Instrumentation and Measurement, Vol. 51, No. 2, April 2002, pp: 235-240

Maria P. F, Santini S, Langthaler P, “Torque generation model for Diesel Engine”, Proceedings of 42nd IEEE Conference on Decision and Control, Maui, Hawaii, USA, Dec 2003, pp: 1171-1176

Matteo M, Nicolo S, “Multiple Misfire Identification by a Wavelet-Based-Analysis of Crankshaft Speed Fluctuations”, IEEE International Symposium on Signal Processing and Information Technology, 2006, pp: 144-148.

Mauer G, “On-line Cylinder Fault Diagnostics for Internal Combustion Engines”, IEEE Transactions on Industrial Electronics, Vol. 37. No. 3, June 1990, pp: 221-226.

Merkisz J, Bogus P, Grzeszczyk R, “Overview of Engine Misfire Detection Methods used in On-Board Diagnostics”, Journal of Kones Combustion Engines, Vol 8, Number 1-2, 2001, pp: 326-341.

Messai N, Thomas P, Lefebvre D, Riera B., Elmoudni A, “Modeling and Monitoring of Hybrid Dynamic Systems with Feed forward Neural Networks: Application to Two Tank Hydraulic System”. Workshop on Advanced Control and Diagnosis, Mulhouse France, 2005, pp: 103-109.

Minghui K, Moskwa J. J, “Model-Based Engine Fault Detection using Cylinder Pressure Estimates from Nonlinear Observer”, Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, FL- Dec. 1994. pp-2742-2747.

Mogens B, Kinnaert M, Lunze J, Staroswiecki M, “Diagnosis and Fault Tolerant Control”, Springer Verlag Berlin Heidelberg 2006, ISBN-10 3-540-35652-5

Mohammad Ali Jinnah University, “Project Funding Agreement, Information System for Early Fault Warning in Automotives” No. ICTRDF/TR&D/2008/10, June 2009.

Monterrubio J. M, Linares R. C, “Sliding Mode Observer for Internal Combustion Engine Misfire Detection”, Fourth Congress of Electronics, Robotics and Automotive Mechanics CERMA 2007 IEEE, pp: 620-624.

Morgan I., H. Liu, “Predicting Future States With *n*-Dimensional Markov Chains for Fault Diagnosis, IEEE Transaction on Industrial Electronics, Vol. 56, No. 5, pp, 1774-1781, May 2009.

Mosterman P. J. ,Eric-J. Manders and Biswas G., “Qualitative Dynamic Behavior of Physical System Models With Algebraic Loops”, Presented at the Eleventh International Workshop on Principles of Diagnosis, Mexico. June 2000

Murphey Y. L, Crossman J. A, Chen Z. H, Cardilo J., “Automotive Fault Diagnosis- Part II: A Distributed Agent Diagnosis System”, IEEE Transactions on Vehicular Technology, Vol 52, No. 4, July 2003, pp:1076-1098

Naik S, “Advanced Misfire Detection using Adaptive Signal Processing”, International Journal of Adaptive Control and Signal Processing, Volume 18, Issue 2, Special Issue: Automotive Applications, 2004, pp: 181-198.

Narendra, K.S. and Parthasarthy, K., (1990), “Identification and Control of Dynamical Systems Using Neural Networks, ” IEEE Trans. On Neural Networks, Vol.1, No.1, pp.4-27

Nyberg M. and Frisk E, “Residual Generation for Fault Diagnosis of Systems Described by Linear Differential-Algebraic Equations”, IEEE Transactions on Automatic Control, Vol. 51, No. 12, Dec. 2006, pp: 1995-2000

Palade V., Cosmin Danut Bocaniala, Lakhmi Jain, “Computational Intelligence in Fault Diagnosis”, Springer-Verlag London Limited 2006, ISBN-10: 1-84628-343-4

Patton, R.J., (1991), “Fault Detection and Diagnosis in Aerospace Systems using Analytical Redundancy, ” Computing & Control Engineering Journal, Vol.2, Issue:3, pp.127-136

Patton R. J. , Uppal F. J, Lopez-Toribio C J, “Soft Computing Approaches to Fault Diagnosis for Dynamic Systems: A Survey”, Proc. of 4th IFAC Symposium SAFEPROCESS, 2000, Budapest, 1:298–311, 2000.

Papoulis A, Pillai S. U, “Probability, Random Variables and Stochastic Processes”, Tata McGraw-Hill, 11th Edition, 2006, pp: 695

Pisu P., Serrani A., You S., Jalics L. “Adaptive Threshold Based Diagnostics for Steer-By-Wire Systems”, Journal of Dynamic Systems, Measurement and Control, Transactions of ASME, June 2006, pp: 428-435.

Pradhan D. K., “Fault Tolerant Computer System Design”, Prentice Hall, ISBN No. 0-13-057887-8, 1995, pp: 55-57.

Protocol document ISO 9141-2, Road vehicles- Diagnostic systems-Requirements for interchange of digital information, Copyright 2000 International Organization For Standardization, Information Handling Services, 2000

Protocol document SAE J1979, Surface Vehicle Standard, rev. 2002, copyrights, Society of Automotive Engineers

Pulkrabec W. W., “Engineering Fundamentals of the Internal Combustion Engines”, Prentice Hall 1997, ISBN: 0135708540

Rizvi M. A., A. I. Bhatti, “Hybrid Model for Early Detection of Misfire Fault in SI Engines”, IEEE 13th International Multitopic Conference, 538315, pp 1-6, Nov. 2009.

Rizvi M. A., A. I. Bhatti, Q. R. Butt “Fault Detection in a class of Hybrid Systems”, IEEE 13th International Conference on Emerging Technologies (ICET) 2009, pp 130-135, Oct. 2009.

Rizvi M. A, Bhatti A. I., Butt Q. R, “Hybrid Model of Gasoline Engine for Misfire Detection”, IEEE Transactions on Industrial Electronics” Accepted for publication in 2010.

Rizvi M. A, Bhatti A. I. , Malik F., Butt Q. R., “Hybrid Model Application for Fault Detection and Quality Assurance”, 16th International Conference on Soft Computing MENDEL 2010, Brno, Czech Republic, ISSN 1803 1803-3814, pp: 93-100.

Rizvi M. A, Bhatti A. I. , Malik F., Butt Q. R., “Misfire Detection in IC Engines using Finite State Automata”, 15th International Conference on Soft Computing MENDEL 2010, Brno, Czech Republic, ISSN 1803 1803-3814, pp: 93-100.

Rizzoni G. , Ribbens W. B. “Crankshaft position measurement for engine testing, control and diagnosis”, Proceedings of the 39th Vehicular Technology Conference, May 1989. Pp: 423-436

Rizzoni, G. Pipe J. G, Riggins R. N, VanOyen M. P. “Fault Isolation and Analysis for IC Engine On-Board Diagnostics”, IEEE, 1988 pp: 237-244.

Rizzoni G. “A Passenger Vehicle On-board Computer System for Engine Fault Diagnosis, Performance Measurement and Control”, Proceedings of the 37th IEEE Vehicle Technology Conference, Tampa, FL, Institute of Electrical and Electronic Engineers, 1987, Page. 450-457

Roemer M. J, Byington C. S, Kacprzynski G. J, Vachtsevanos G, “An Overview of Selected Prognostic Technologies with Application to Engine Health Management”, Proceedings of GT2006 ASME Turbo Expo 2006: Power for Land, Sea, and Air May 8-11, 2006, Barcelona, Spain, pp: 707-715

Schutter D, Heemels W, Bemporad A, : “Modeling and Control of Hybrid Systems”, Lecture Notes of the DISC Course, Delft Center of Systems and Control, Delft University of Technology, 2003, pp. 67-82

Smith F. S, Shen Q, “Fault Identification Through the Combination of Symbolic Conflict Recognition and Markov Chain Aid Belief Revision”, IEEE Transactions on Systems, Man and Cybernetics- Part A: Systems and Humans, Vol-34, No. 5, Sept 2004. pp: 649-663.

Sood A. K, Fahs A. A , Naeim A. Henein, “Engine Fault Analysis: Part II- Parameter Estimation Approach”, IEEE Transactions on Industrial Electronics, Vol IE-32, No. 4, Nov. 1985, Pages: 301-307

Sood A. K. , Friedlander C. B, Fahs A. A., “Engine Fault Analysis: Part I-Statistical Methods”, IEEE Transactions on Industrial Electronics, Vol. IE-32, No. 4. pp. 294-300, Nov 1985,

Speyer J. I, Chung W. H, “Stochastic Processes, Estimation and Control”, 1st Edition. 2008 SIAM, pp. 157-159

Stone. R, Jeffery K. Ball, “Automotive Engineering Fundamentals”, SAE, ISBN 0-7680-0987-1, 2002

Stotsky A. A, “Statistical Engine Misfire Detection”, Proceedings of the Institution of Mechanical Engineers. Part D, Journal of Automobile Engineering ISSN 0954-4070, Volume 221, 2007, pp: 641-649.

Suchomski P, “High-order Interacting Multiple Model Estimation for Hybrid Systems with Markovian switching parameters”, International Journal of Systems Science, 2001, vol 32, No. 5, pp: 669-679.

Sun S, Yu G, Zhang C, “Short Term Traffic Flow Forecasting using Sampling Markov Chain Method with Incomplete Data”, IEEE Intelligent Vehicles Symposium, University of Parma, Italy, June 2004, pp: 437-441

Swets, J., “Measuring the accuracy of diagnostic systems”. Science 240, 1988, pp: 1285–1293.

Tabuada P., “Verification and Control of Hybrid Systems”, Springer, Dordrecht, Heidelberg, London, NewYork, 2009, ISBN: 978-1-4419-0223-8.

Tinaut F. V, Melgar A, Laget H, Dominguez J. I, “Misfire and Compression Fault Detection through the Energy Model”, Mechanical Systems and Signal Processing, Volume 21, Issue 3, April 2007, pp: 1521-1535.

Torrisi F. D, Bemporad A, “HYSDEL-A Tool for Generating Computational Hybrid Models for Analysis and Synthesis Problems”, Transactions of Control Systems Technology, Vol 12, No. 2, March 2004. pp. 235-249

Uppal F. J, Patton R. J, “Fault Diagnosis of an Electro-pneumatic Valve Actuator Using Neural Networks With Fuzzy Capabilities”, ESANN'2002 proceedings - European Symposium on Artificial Neural Networks, Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 501-506

Vemuri A. T. “Sensor Bias Fault Diagnosis in a Class of Nonlinear Systems”, IEEE Transactions on Automatic Control, Volume 46, Number 6, 2001, pp: 949-954

Villarino R, Bohme J. F, “Pressure Reconstruction and Misfire Detection from Multichannel Structure Borne Sound” ICASSP 2004, pp:II-141-II-144.

Walter A, Kiencke U, Jones S, Winkler T, "Misfire Detection for Vehicles with Dual Mass Flywheel (DMF) Based on Reconstructed Engine Torque" Document Number: 2007-01-3544, 2007, University of Karlsruhe.

Weeks W. W, Moskwa J. J, "Automotive Engine Modeling for Real Time Control Using Simulink" SAE 950417, 1995, pp: 1-15.

Willskey, A.S., (1976), "A Survey of Design Methods for Failure Detection in Dynamic Systems, " Automatica, Vol.12, pp. 601-611

Witczak M., "Advances in Model Based Fault Diagnosis with Evolutionary Algorithms and Neural Networks.", International Journal of Applied Maths Computer Science, 2006, Vol. 16, No. 1, 85–99

Wong P. K, Tam L. M., Li K, Vong C. M, "Engine Idle Speed System Modeling and Control Optimization using Artificial intelligence", Proceedings of the Institution of Mechanical Engineers. Part D, Journal of Automobile Engineering, Volume 224, 2009, pp: 55-72.

Wu, N.E., (1992), " Failure sensitizing reconfigurable control design, " Proc. of the 31st IEEE Conf. On Control & Decision, Tucson, AZ, pp. 44-49

Wu Z. J, Lee A, "Misfire Detection Using a Dynamic Neural Network with Output Feedback", Document Number: 980515, 1998

Wünnenberg,J., (1990), "Observer-based Fault Detection in Dynamic Systems," PhD thesis, Univ. of Duisburg, Germany

Yaojung S, Moskwa J. J, "An Investigation of Load Force and Dynamic Error Magnitude using the Lumped Mass Connecting Rod Model", SAE Technical paper 930617, 1993.

Yan L, Miller D. J, "General Statistical Inference for Discrete and Mixed Spaces by an Approximate Application of the Maximum Entropy Principle", IEEE Transactions on Neural Networks, Vol II, Number 3, May 2000, pp: 558-573

Ying H., Petros A. Ioannou, Majdedin Mirmirani, "Fault-Tolerant Control and Reconfiguration for High Performance Aircraft: Review", University of Southern California Dept. of Electrical Engineering-Systems.

Yu X. G, Liu Y. H., Wei D, Ting M. "Hybrid Markov Models Used for Path Prediction", IEEE 15th International Conference on Computers Communications and Networks ICCN, 2006, pp: 374-379.

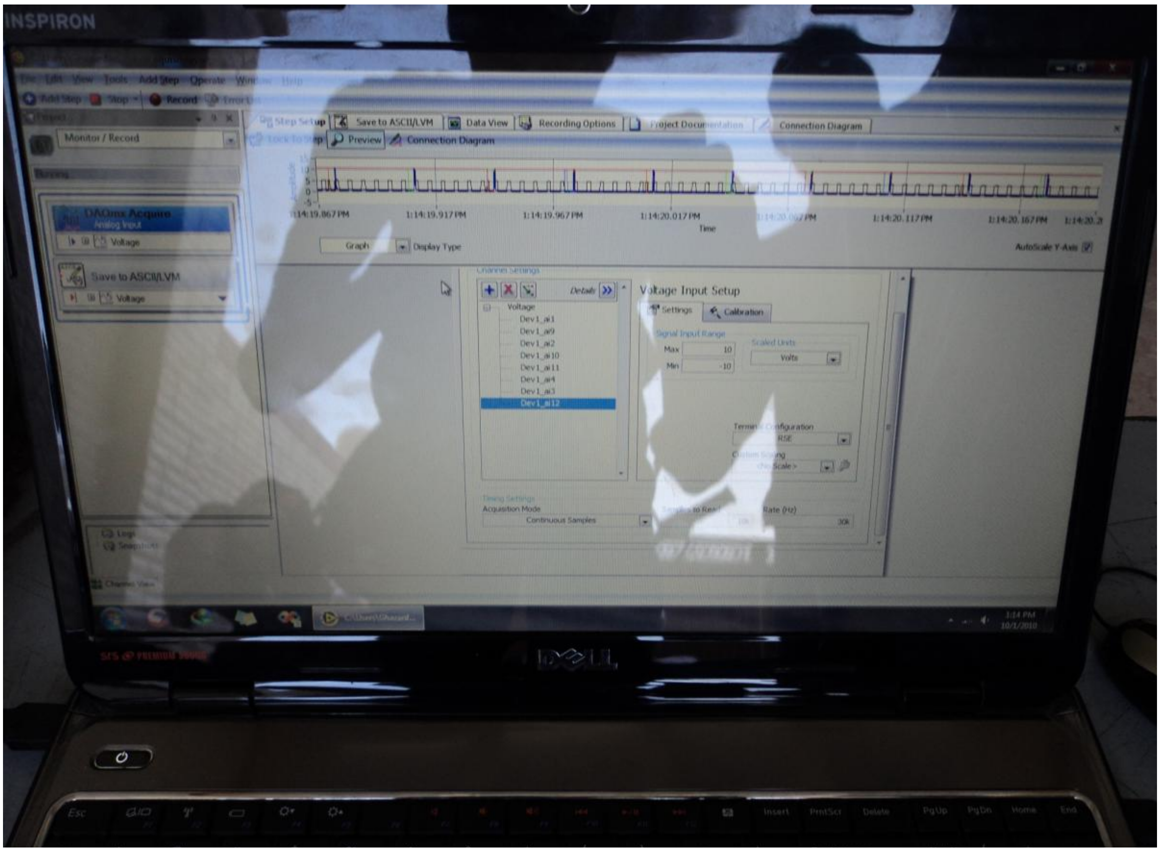
Yu M, Luo M, Wang D, Arogeti S, Zhang X, "Simultaneous Fault and Mode Switching Identification for Hybrid Systems based on Particular Swarm Optimization", Expert Systems with Applications 37, (2010), pp: 3000-3012

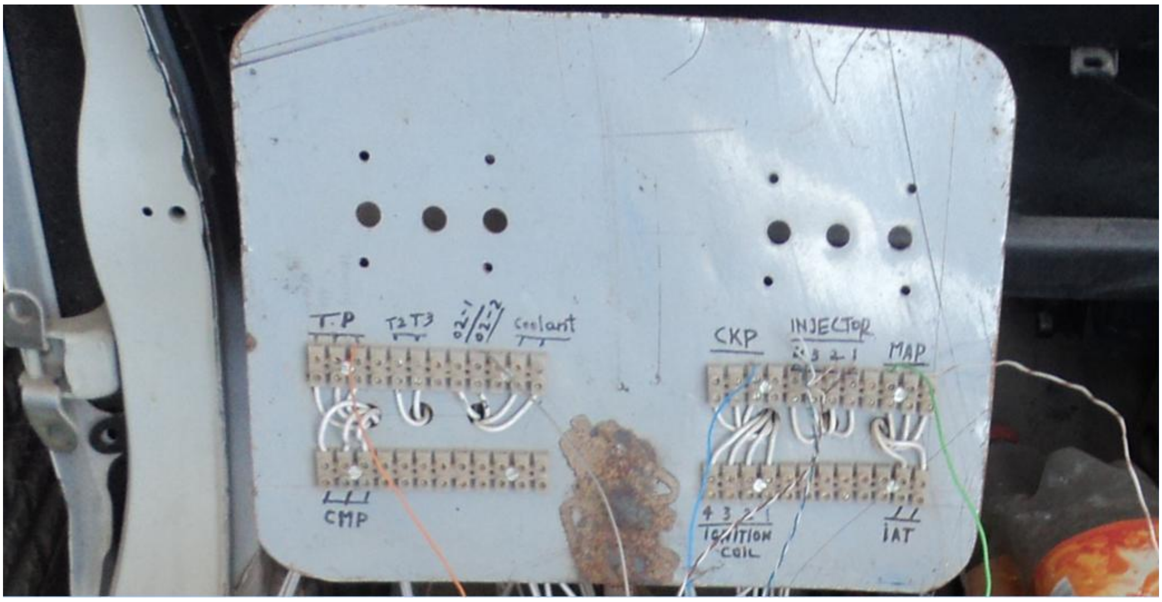
Zhendong S, Shuzhi S. G, "Switched Linear Systems", © Springer Verlag London Limited 2005, ISBN 1-85233-893-8, pp: 16-18

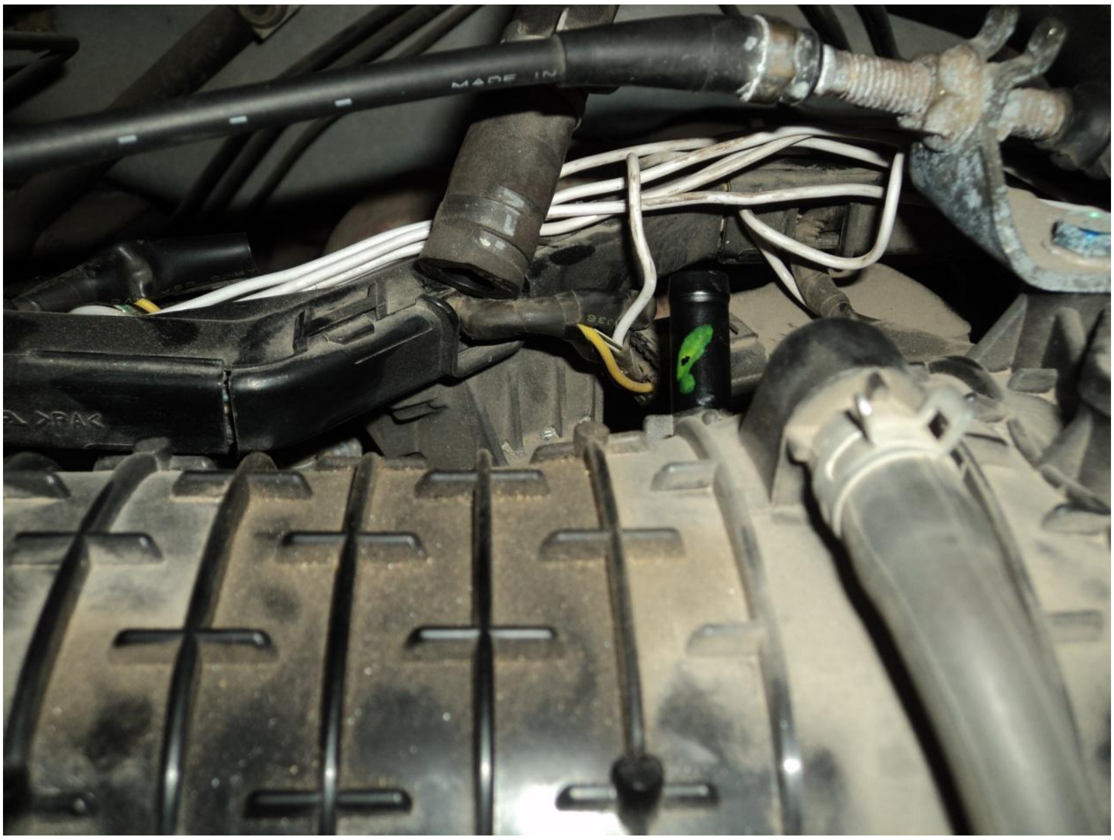
Zhinong L, Zhaotong W, Yongyong H, Chu F, "Hidden Markov Model Based Fault Diagnostics Method in Speed up and Speed down Process for Rotating Machines", Mechanical Systems and Signal Processing, Volume 19, 2005, ISSN: 0888-3270, pp: 329-339.

APPENDICES

APPENDIX - A: Experimental Setup




T.P
T2 T3
Coolant
CKP
INJECTOR
MAP
CMP
IGNITION
COIL
IAT


MADE IN

APPENDIX - B: Typical Data Of SI Engine

Typical data of some automotive engines provided by Srinivasan

Vehicle Name	FordEscort 15V	Opel Astra 1.6	Maruti Esteem VX
Engine Bore	76 mm	79 mm	74 mm
Engine Stroke	78 mm	81.5 mm	75.5 mm
Engine Capacity	1597 cc	1598 cc	1298 cc
Engine Compression ratio	8.8 : 1	9.2 : 1	9.2 : 1
Max. Power at speed (bhp)	83	76	65
	5500 rpm	4500 rpm	6000 rpm
Max Speed (rpm)	7200	6000	6500
Max. Torque (N-m) at speed	124	121	99
	3000	2500	4000