Military Technical College
Kobry El-Kobbah,
Cairo, Egypt

10th International Conference
on Electrical Engineering
ICEENG 2016

Design and Implementation of Algorithm of Wireless Communication Data Frame Recognition

By

ZHANG Yi-jia *
LIU Rui-ying **

Abstract:

This paper analyses several typical data-frames of wireless communication, proposes an algorithm of data-frame recognition, which is based on tagged word of data-frames, and analyses the performance of algorithm. This paper resolves problems of protocol type analyzing and data frame recognizing, which are for date link layer (DLL) of wireless communication, by the algorithm of data-frame recognition.

Keywords:

Recognition, data-frame, false acceptance rate, false rejection rate, tagged word

* 1. Science and Technology on Communication Information Security Control Laboratory, Jiaxing, China
2. Jiangnan Electronic Communication Research Institute, Jiaxing, China
17081912@qq.com

** College of Mechanical and Electrical Engineering, Jiaxing University,Jiaxing, China, 18266632686@163.com

1. Introduction:

Along with the popularization and application of network, the network protocols also emerge in endlessly. Data can only be transmitted on the network after being packaged as packets based on certain agreement format. In the aspect of network confrontation, especially the wireless network, data are mainly obtained through the interception of the link of wireless network. While protocol type and data link layer frame can't be directly obtained from network devices, so the analysis of the protocol and the identification of data frame must be done for acquiring these parameters. At present, the analysis technology for upper level protocol is very popular, some of the most famous software and hardware are ethereal, Sniffer, airpcap and so on [1]. But for there is almost no analysis algorithm or product for bit information obtained through demodulation and decoding directly from the physical layer. Therefore the protocol analysis and data frame identification algorithm is one of the most important ways to improve the ability of electronic warfare on battlefield information network [2].

In this paper, the protocol analysis for DLL is to achieve filtering and identification on bit information of DLL which is intercepted from the network, and then to discriminate protocol type. The identification of data frame is also for filtering and recognition of the intercepted bit information, getting a different type of protocol packet, then obtaining the upper protocol type through the further interpretation of packet. Hence the recognition of the data frame on link layer is the basis of the parsing of upper protocol [3].

The analysis of DLL protocol and identification of data frame on wireless communication are mainly studied in this paper; three typical wireless communication systems such as AIS and ACARS, WLAN are selected for the analysis. In this paper, the data acquisition method is introduced firstly, which is the basis of identification; Then the analysis research is done on the data frames of the three typical protocols; a recognition algorithm of data frame is put forward; the performance of the algorithm analysis, experimental results and conclusions are given at last.

2. Design and realization of the algorithm:

A. Recognition Basis of Data Frame

Wireless network is the collection constituted by many independent communication nodes which are linked with each other through wireless communication channels; its basic elements can be divided into network nodes and communications channels. The basis of identification data on wireless DLL is to demodulate data successfully and obtain bit stream by channel decoding. In this paper, the analysis and processing are only done on DLL, without considering the

demodulation and decoding of physical layer and the real-time analysis and processing of data.

This paper chooses AIS, ACARS, WLAN for wireless communication protocol analysis research. Through the research, tagged words of protocol can be extracted as many as possible to combine with other feature words for protocol identification, which can lay the foundation for the fundamental and algorithm of data frame recognition.

(1) The main function of shipborne's AIS is to broadcast the identification information, location information, movement parameters, sailing state and other important data related to navigation safety to the surrounding ship through Very High Frequency (VHF) signal in order to realize the identification and monitoring of shipping [4]. The DDL of AIS applies high-level data link control (HDLC) protocol; the message structure, sequence and the relationship between data grouping and time slot are shown in figure 2. The total length of a packet is 256 bits. The head of 24 bits make up with alternating synchronous sequence of 0 and 1. Synchronous sequence is applied to train the receiver to accept synchronous data bit; the following bit is the start tag (0x7e), which indicates the head of the packet. 16 bits of cyclic redundancy check (CRC) is for the error detection of package [5]. Then the end tag follows detection data, it is the same with the start tag. The bit stream must be filled according to the following method during data transmission, once there appears more than five ones in the output bit stream, 0 should be inserted into the bit stream, this method is applied to any bit except the HDLC tag data. The end 24 bits of the packet is buffer period including some content: bit filling, distance delay, repeater delay, synchronous shaking.

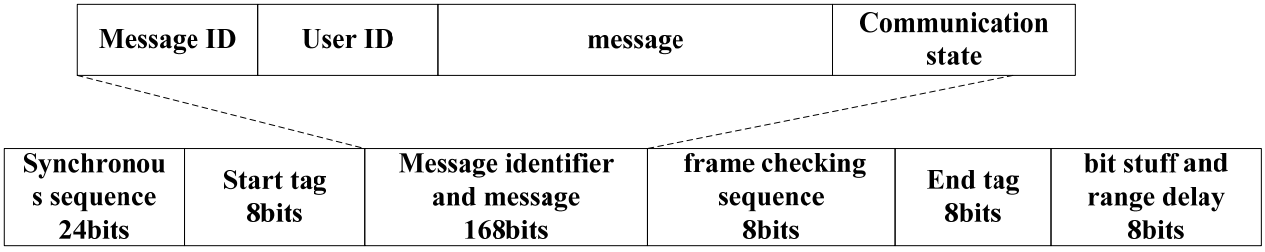


***Figure (1):** Packet format of AIS data frame*

(2) Air-ground data link systems based on Aircraft Communication Addressing and Reporting System (ACARS) are commonly used in civil aviation field [6]. The system adopts the VHF Data Link MODE 1 (VDL1) for information transmission, the frequency of ACARS is 118MHz-137MHz, data transfer rate of air ground is 2.4Kbps, and channel spacing is 25KHz. VDL1 based on ACARS is encapsulated through the ARINC-618. ARINC-618 is a character-oriented protocol between aircraft and ground station, and also the communication protocol between aircraft and chain providers during flight. Table 1 shows the message format of ARINC-618 [7].

***Table (1):** The message format of arinc-618*

Tagged word	Bytes number	Tagged word	Bytes number
SOH	1	MSN	4
Mode	1	FlightID	6
Address	7	AppText	0-220
TAK	1	Suffix	1
Label	2	BCS	2
DBI	1	BCSSuffix	1
STX	1		

Here, SOH (0x01), mode, address, TAK, Label and DBI/UBI make up the message header to control the transmission of message; STX (0x02) is the beginning tag of content, Suffix (0x03) is the end tag of content. While the message content are between STX and Suffix; BCS is the CRC check code for the message, BCSSuffix (0x7f) is the end of the message. The format of message header field is fixed and involved in communication transmission, while the message content is determined by message type and the state of aircraft [8].

(3) In recent years, the wireless local area network (WLAN) which sets 802.11 series as standard protocol has obtained great development. 802.11a owns a mainstream standard working frequency band at the 5GHz, and support the transmission rate of 54 Mbps through applying OFDM modulation technology. Figure3 shows the format for the PPDU including the DSSS PLCP preamble, the DSSS PLCP header, and the MPDU. The PLCP Preamble contains the following fields: Synchronization(Sync) and Start Frame Delimiter (SFD). The PLCP Header contains the following fields: IEEE 802.11 Signaling (Signal), IEEE 802.11 Service (Service), LENGTH (Length), and CCITT CRC-16 [9].

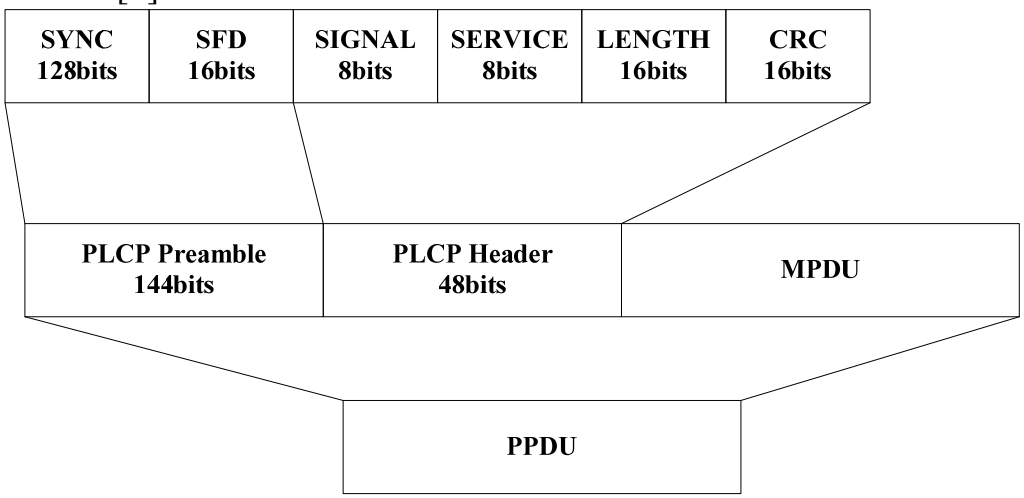


***Figure (2):** PLCP frame format*

B. The Analysis of The Data Frame Structure

Through the analysis of data frame structure on the three protocols, the defined bits can be acquired as shown in table (1). The so-called tagged words are the fixed bits of data frames with the same protocol type under any possible conditions. The value of tagged words is fixed, for different application environments; the same protocol still has the same tagged words. There are positive numbers and negative numbers in table (2). Here positive numbers represent the head bits relative to the data frame, negative numbers represent the tail bits relative to the data frame. These feature bits can be applied to filter and identify data frames of different protocols, to locate each frame position in the data flow and analyze the content of the data field, then to obtain the upper layer protocol type.

***Table (2):** Feature position of data frame*

protocol	Feature position(sequence number and numerical value)				
AIS	Location	1 to 8	9 to 16	17 to 24	
	Tagged words	01010101	01010101	01010101	
	Location	25 to 32	33 to 40	-32 to -25	
	Tagged words	01010101	01111110	10101010	
ACARS（uplink）	Location	1 to 8	105 to 113	-32 to -24	-1 to -8
	Tagged words	00000001	00000010	00000011	01111111
WLAN (802.11a)	Location	1to128		129 to 144	
	Tagged words	1		0	

C. The Identification Algorithm of Data Frame

An identification algorithm of data frame called container algorithm is proposed in this paper. The flow of algorithm is shown in figure (3).

So-called container algorithm is to put the special bits of the listened binary data stream into several certain cache blocks, these cache blocks are called containers. Then we will match value of several present fixed feature bits with the cache content in the container, if they are the same, the protocol frame will be filtered out, and also the protocol counter will be added with a minimum frame length. Then the container algorithm will be again applied to do identification after skipping the data frame header. However, if they are different, a bit data will be skipped to continue to match. Here, three processes of AIS, ACARS and WLAN are independent with each other, without affecting each other in the implementation process, this means that every piece of data need to undergo the container algorithm for three times.

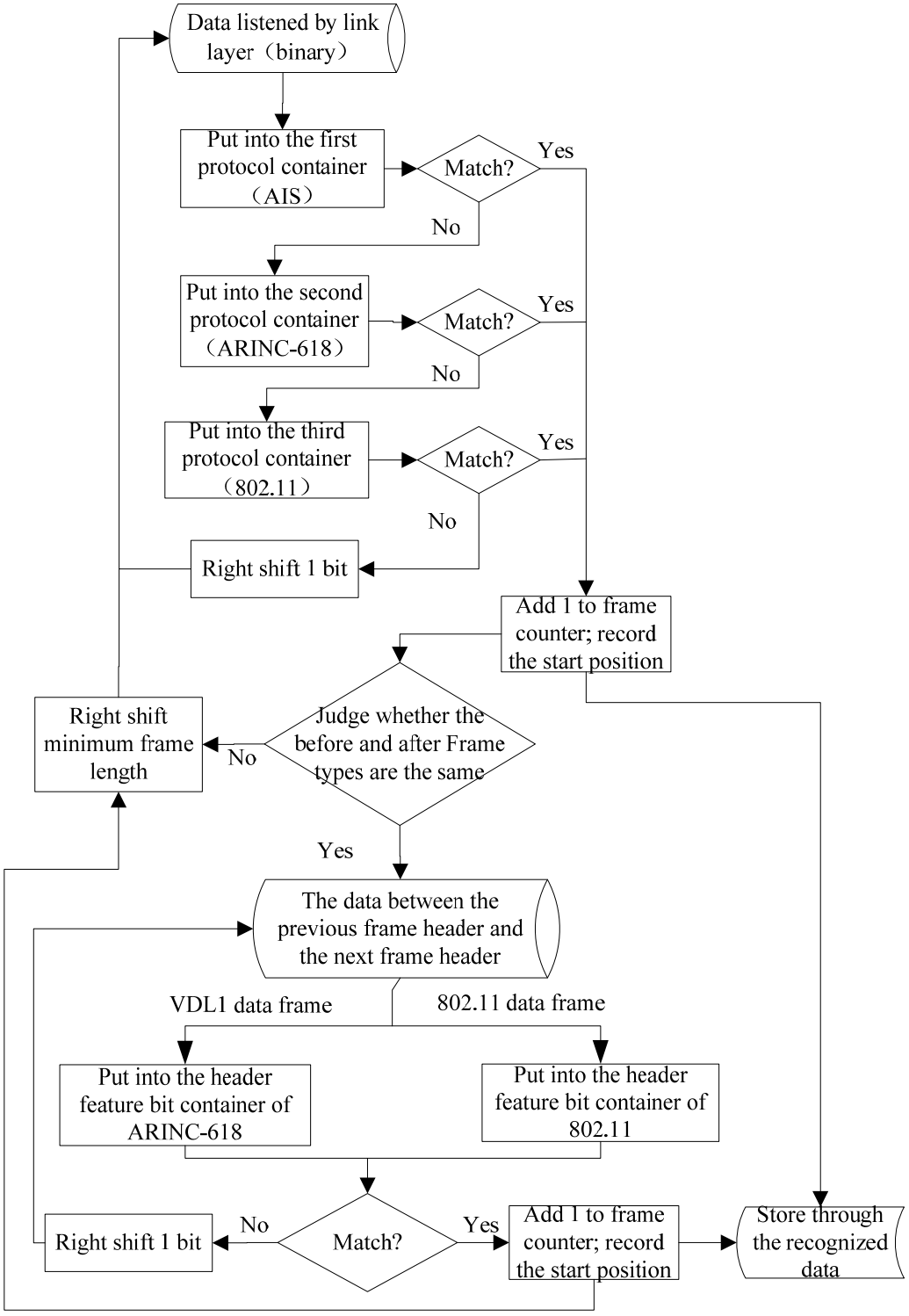


***Figure (3):** The flow chart of container algorithm*

3. The performance analysis and simulation experiments：

As for the identification requirements of wireless communication network protocols, we only need to identify one or several frames in the data, which can be enough to judge out the DDL protocol type and the upper protocol type. Hence there is no need to do identification and analysis for all the data frames. And when evaluating the performance of identification algorithm, we mainly consider the false acceptance rate of data frame recognition, whereas secondly consider the false rejection rate of data frame recognition.

A. Performance Analysis

The occurrence probability of 0 to 1 in any random bit of listened binary data is 0.5, and every bit is independent with others, which meets the specific conditions of classical probability (limitation, equal-possibility and exclusiveness). Hence false acceptance rate of protocol frame can be calculated through multiplication principle of classical probability. The specific calculation is shown in table (3).

***Table (3):** Acceptable false rate of container algorithm*

System	Protocol	Acceptable false rate
AIS	HDLC	$P(\text{AIS}) = 0.5^{48} = 3.553 \times 10^{-15}$
ACRAS	ARINC-618	$P(ACRAS) = 0.5^{32} = 2.328 \times 10^{-10}$
WLAN	802.11a	$P(802.11) = 0.5^{144} = 4.484 \times 10^{-44}$

From the probabilities in the table (3), as the number of tagged words about 802.11a protocol is the biggest, the calculated false acceptance rate is the smallest, which can basically meet the requirements of protocol identification. There is no need to modify the identification algorithm. However the tagged words of ARINC-618 and HDLC are less than 802.11a, the calculated false acceptance rate is much higher, so the container algorithm must be improved to meet the practical requirements.

B. The Improvement of Container Algorithm

In the wireless communication process of the network, the data frames are correlative. In the case, data frames of the same agreement keep continuous sending. Hence we can adopt the related identification method to modify the original algorithm in order to reduce false recognition rate of data frame.

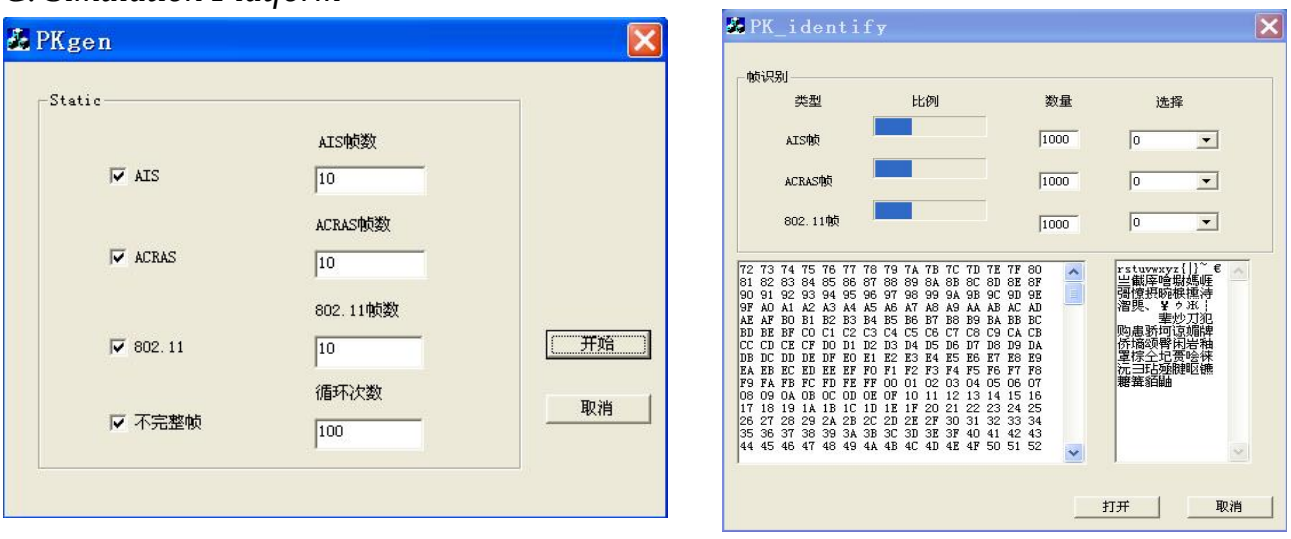
The so-called related identification method is to judge whether the before and after identified types of data frames are the same or not through the characteristics of the before and after data frame correlation. If they own the same type, the data frame will be considered as a true one; if not, the container algorithm for matching single data frame will be applied to do identification. The related identification method can greatly reduce

the false acceptance rate of data frame, which can be calculated as shown in table (4). Here, *n* represents the number of the joint-matching data frames. However this method also has its drawbacks; the false rejection rate will be increased when the false acceptance rate is decreased. Especially when data frames with several different types need to be transmitted alternately and frequently, there will be a quite high false rejection rate. Theoretically, the bigger the value of n is, the lower the false acceptance rate is. But when there are different link layer frames in the wireless network, if the value of n is too big, false rejection rate of the algorithm will exceed the permissible limits. Therefore the value of n must be reasonably selected considering about actual situation.

***Table (4):** Acceptable false rate of modified container algorithm*

System	Protocol	Acceptable false rate
AIS	HDLC	$P(AIS)=(0.5^{48})^n=(3.553\times10^{-15})^n$
ACRAS	ARINC-618	$P(ACRAS)=(0.5^{32})^n=(3.328\times10^{-10})^n$
WLAN	802.11a	$P(802.11)=(0.5^{144})^n=(4.484\times10^{-44})^n$

C. Simulation Platform



(a) Generating program of data frame (b) Identification program of data frame

***Figure (4):** Simulation software*

The corresponding simulation experiment software is established for testing the proposed containers algorithm and modified algorithm. This software is made up with the generating program and identification program of data frame and is coded by Visual

C++. The interface is shown in figure (4). PKgen is applied to generate data frames. The simulation data of data frame will be generated after entering the AIS frame number, ACRAS frame number, WLAN frame number, their cycles and clicking on the "start" button. PK identification is applied to identify the data frame which can obtain the corresponding frames and proportion of the given frames after inputting the generated simulation data.

D. Simulation Results

Generating program is used for generating data with different lengths and different protocol types; and identification program is used for identifying the generated data. The simulation results are shown in table (5). From the simulation results, the results conform the above probability calculation, while the number of frames is larger; the false acceptance rate of data frames with AIS and ACRAS will be larger than the threshold. It will cause that the data section would take mistake to be a tagged word. At this time the modified container algorithm can be adopted in order to reduce the false acceptance rate.

In wireless network, the algorithm is only used to do identification and parsing on DLL protocol. Assuming the message transmission rate is 10 Mbps, the average length of data frame is 5Kbits, that is 2000 frames can be obtained only in about 1 second, the data stream can basically include all types of data frames in DLL, so the container algorithm can be used for the discriminant analysis of DLL's protocol type, and the modified container algorithm can be used to identify data frames.

***Table (5):** Partial test results*

protocol type / false acceptance frame number / total frame number	AIS	ACRAS	WLAN
100000	0	0	0
900000	6	0	0
3000000	56	24	0
6000000	176	26	0

4. Conclusions:

This paper proposes a practical container algorithm and the modified algorithm for identification of data frame through the frame structure analysis on three typical DLL protocols in the wireless communications. The performance analysis and simulation

results are given at last. The results show that container algorithm owns a lower false acceptance rate and can be applied in protocol analysis and data frame identification from the wireless communications. Container algorithm is not only suitable for protocol analysis and identification in wireless communication, but also can be extended to the identification of the wire net protocols, wire communication protocols, and other various protocols as it has the feature of strong modularity and universality.

References:

[1] A. Willig, K. Matheus and A. Wolisz. *Wireless Technology in Industrial Networks [J]*. Proceeding of the IEEE, 2005, 93(6): 1130-1151.

[2] R. Milner. *Comunicaton and concurrency [M]*. California, U.S.A: Englewood Cliffs, 1999. 463-467.

[3] D. B. Faria and D. R. Cheriton. Detecting identitybased attacks in wireless networks using signalprints. In Proc. of *ACM Workshop on Wireless Security*, September 2006.

[4] B J Tetreult , "Use of Automatic Identificaiton System (AIS) for maritime domain awareness (MDA)[A]", *MTS/IEEE. OCEANS.2005*. Proceedings of MTS/IEEE [C] , pp.1590 -1594

[5] Hu Yueli , "Data Link Capacitance and Congestion Resolution for AIS" , *Computer Measurement & Control* , pp.1631 -1634 , 2007

[6] ICAO. *Doc 9694-AN/955 Manual of Air Traffic Services Data Link Applications*, First edition, 1999.

[7] EUROCONTROL. *Introduction to VDL2 Capacity Analysis through Simulations. Analysis of Link2000+ Initial Deployment (Single Channel).* Edition 2. 3, January 2006.

[8] Bretmersky, Steve, et al., 2005, *Characteristics and Capacity of VDL Mode 2, 3, and 4 Subnetworks, Journal of Aerospace Computing, Information, and Communication* Vol. 2, pp. 470-489.

[9] IEEE 802.11: *Wireless lan medium access control and physical layer specifications*. IEEE Computer Society LAN MAN Standards Committee, August 1999.

Notes:

ZHANG Yijia(1981-), he received his B.S. and M.S. degrees from Harbin Institute of Technology in 2006. Now he is a PhD Candidate in China University of Petroleum. His research field focuses on electronic warfare.

Liu Ruiying(1990-), she received her B.S. degree from China University of Petroleum in 2011. Her research field focuses on Electrical system design and Fault detection and diagnosis.