Intelligent diagnostic for sensor arrays

Elena Gaura, Robert M. Newman

Coventry University, Priory St, CV1 5FB, UK
e.gaura@coventry.ac.uk

ABSTRACT

The paper discusses the appropriateness of local sensor health monitoring/sensor fault diagnosis and suggests an Artificial Intelligence (AI) based solution to the diagnosis of sensor faults, in the context of sensor arrays. In contradistinction with the "centralized" methods currently used process/system fault detection, the work presented here aims to establish a process independent and sensor specific methodology for detecting and isolating sensor failures and fault recovery. The intelligent system proposed integrates the adaptability and learning ability of Artificial Neural Networks with effective inter-sensor communication protocols. Micromachined accelerometers are considered as a case study.

Keywords: Sensors, fault diagnosis, artificial intelligence, neural networks.

1 INTRODUCTION

At present, the requirements for safe and reliable operation of processes and systems extend beyond the normally accepted safety-critical systems (e.g. nuclear reactors, chemical plants, and aircraft), to new systems such as autonomous vehicles and rapid transport systems. Early detection of faults and/or malfunctions in industrial processes and systems can help reduce downtimes and the incidence of catastrophic events. As sensors are essential components of any process or system which makes use of automatic control, it follows that an important aspect of any process/system fault diagnosis strategy is to be able to assess their state of functionality. Whilst the common approach to sensor health checks is currently through periodical calibration against a given "ideal" sensor, more and more applications are being developed where such checks are either impossible or insufficient. In many manufacturing industries, for example, most sensors are built into equipment and, once the specific line is installed, they are no more directly/easily accessible to operators; fault-finding in this case is difficult and dangerous, and troubleshooting is time-consuming .What is needed therefore, is a system with smart devices that would "speak up" when fault occur.

Considerations such as the above, together with the ever-increasing complexity of present days automated systems lead to new ways of fault detection and fault tolerant control strategies being devised. Most of these techniques, however, are designed to work in a 'centralized' manner, by accounting simultaneously for sensors, actuators and process/system component faults. They are designed specifically for the systems in hand and are not transportable.

Recently, the idea of 'hierarchical' system design with respect to sub-components fault diagnosis has however been claimed to be effective for large systems and, in this context, it could be assumed that there is scope for exploring 'local' diagnosis methods for sensors.

It is intended here, to argue the feasibility of using Artificial Neural Network (ANN) techniques for implementing such localized diagnostic systems through enhancing the individual sensor processing and decision capabilities.

ANNs have been successfully used in a variety of applications for complex data analysis and feature extraction [1]. In the context of the proposed discussion, the associative and predictive nature of an ANN is used for detecting and isolating failures.

Once the ANN is trained for a particular task, operation consists of propagating the data through the mapping produced by the ANN, thereby making possible real-time self-diagnosis, self-validation and monitoring. Acceleration sensors provide a good example for the discussion, as their lack of accessible internal signals makes the tasks of diagnosis and validation particularly challenging.

2 SENSOR SELF DIAGNOSIS

In most applications, accurate and reliable sensor readings are vital for good overall system performance [2]. Despite advances in fabrication technologies, sensors generally exhibit imperfections (for acceleration sensors, for example, common imperfections are: offset, drift, non-linearity and noise) and the magnitude of these imperfections is found to vary both from sensor to sensor and with time. Fundamental characteristics of the sensor, e.g. sensitivity, may be subject to manufacturing tolerances, varying material properties and ambient effects [1]. Moreover, during operation, as with any other system component, sensors may develop several types of faults and fail in a variety of ways.

Over the last few years sensors have evolved into ‘smart’ or ‘intelligent’ versions, where known, fabrication-process inherent sensor imperfections are corrected through post processing of the electrical sensor signals [2].

Equally, the in-work (unpredictable) sensor fault/failure problem are mostly overcome through the use of hardware redundancy or majority voting (this is a typical solution in safety-critical systems, for example, civilian aircraft) [3]. However, such a ‘collective’ measurement validation technique is often prohibitive. For example, one of the problems associated with systems comprising large number of accelerometers, is that of data analysis bottlenecks. The data analysis generally includes: (1) Fourier analysis; and (2) statistical methods for determining outliers and inconsistent measurements, both of which are computationally expensive processes.

Another approach to validate measurements obtained from sensors and to diagnose sensor faults is to use mathematical/knowledge-based modelling of the system under measurement [3], to identify inconsistencies in measurement data. Such a system/process specific approach has two major drawbacks [4]:

- detailed mathematical models of the system/process are required; these are generally extraordinarily complicated to construct and may have significant errors (although some success has been recently reported on using ANN techniques for the modeling and fault detection of such systems, [3]);
- the validation algorithms are ‘tuned’ for each system/process. Retuning, due to a slight modification of the system or the introduction of the same techniques to a new application, is extremely expensive and demands a high degree of expertise.

It follows that, expending research effort on sensor self-validation techniques could be justified.

For acceleration sensors, no definitive results have been found in the literature referring to possible methodologies/implementations of self-validation and diagnosis capabilities within the sensor structure, although some interest in the field was shown by several research centers [2].

In this research, the authors took a systemic approach to the sensor design in order to expand its capabilities. The embedded microprocessor market is growing very fast, and embedded modules reported in the literature are capable of not only computation but also communication; take for example wireless integrated network systems, where it is proposed that systems consist of self contained nodes composed of sensor, actuator, interfaces, data fusion circuitry, general purpose signal processing and microcontrollers.

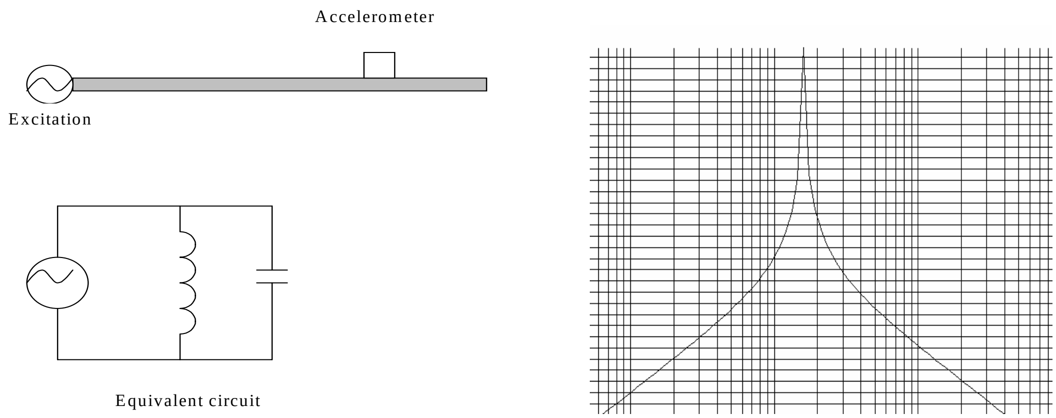


Figure 1: Beam model of structure under test, equivalent circuit and resonance

With modern integrated fabrication technologies it is possible to design and fabricate complete systems in a low cost module, including sensing and powerful signal processing. As well as the analog and digital signal processing required for the function of the sensor itself there can be sufficient capacity to allow the sensor to take on more advanced computational functions, so that system level functionality can be deployed in the sensors themselves.

Such sensors could provide the opportunity to build sophisticated systems consisting of many collaborating sensors, of increased reliability, once the self-validating functions are built-in. This facility could also provide advantages at the level of the overall operation of the industrial systems and processes that the sensor network is part of. Localized, real-time sensor self-validation, self-diagnosis technology will enable improved automatic process/system monitoring and control, removing this overhead from a central process controller. Early detection of small, incipient (rather difficult to detect) sensor faults can be achieved and therefore downtime can be reduced and catastrophes can be avoided. More robust, fault-tolerant control can be designed on the above basis, to accommodate/compensate for soft sensor failures (i.e., recoverable sensor failures that leave no permanent damage). Finally, easy identification of sensors which have suffered hard failures (irrecoverable sensor failures) can be achieved.

The design requirements of such enhanced sensors are to enable unique sensor failure diagnosis and measurement validation with minimal sensor requirements (no hardware redundancy is necessary), by exploiting the information content of readily available signals: the sensor output signal and contextual information gathered from the sensor’s working environment. The above capabilities could be incorporated into a validation and diagnosis module (VDM), which, associated with the sensor, should be able to detect, in real-time, several common sensor faults and failures and issue specific warnings and provide confidence indices for each validated measurement value.

Prerequisites for this design activity are:

- identification (based on experimentally obtained sensor signatures) of features which characterise several common sensor faults.
- determination of the nature of additional, application related information, which can be used in conjunction with the sensor output for fault diagnosis and measurement validation purposes.
- assessment of the feasibility and applicability of using ANN techniques for qualitatively and quantitatively representing the information gathered above.

The remainder of this paper presents an initial analysis of some of these prerequisites, using acceleration sensors as a working example.

3 ACCELERATION SENSORS – EXAMPLE

In the development of a reliable signal-based diagnosis and validation strategy, there is a need to consider not only sensor failures (soft and hard) but also situations which can give rise to faulty (false) measurement data, in the specific application where the sensor is used. For the chosen case study, a set of fault signatures for acceleration sensors could be obtained under laboratory conditions for several different scenarios.

To gain an initial idea of the nature of these signatures different types of sensor fault conditions were modeled, using electrical equivalent circuits.

Case 1. The sensor being not rigidly attached to the structure under test. We can imagine a simplified version of this case as illustrated in Figure 1. The test structure is represented by a beam, which will have some characteristic vibration frequency. The improperly mounted sensor is mounted on this beam by a compliance or spring. This is in itself a resonant system, with some characteristic resonance given proportional to the mass of the sensor and the stiffness of the spring. We can produce an electrical analogue of this system, as shown in Figure 2.

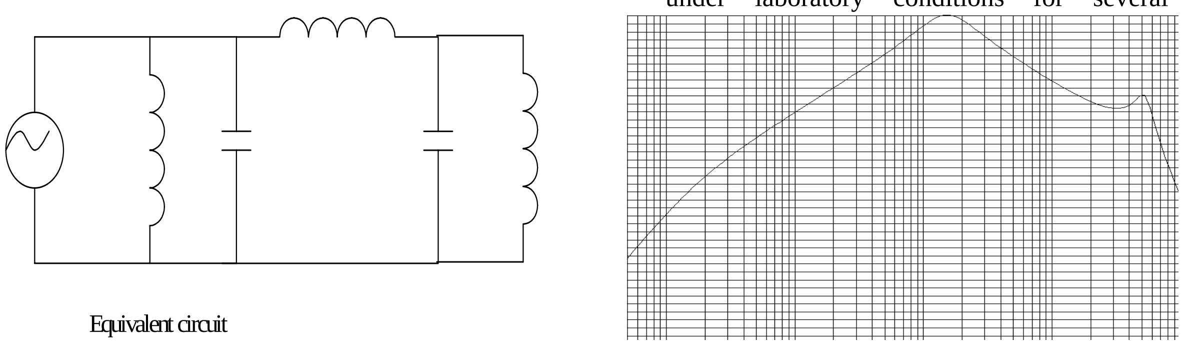


Figure 3: Beam fractured near accelerometer, model and resonance

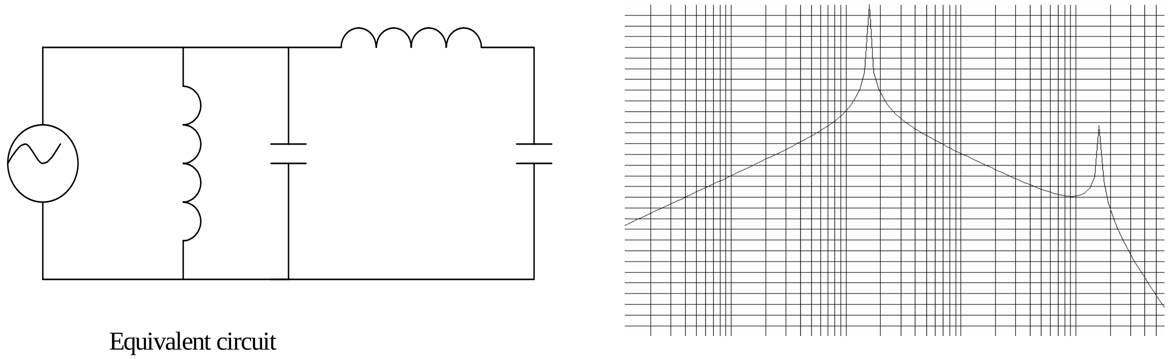


Figure 2: Detached accelerometer model and resonance

So long as the resonant frequency of the poorly mounted sensor is widely separated from that of the beam, then the symptoms of the mounting fault can be clearly differentiated. Since a micromachined sensor is likely to be very light, and even a poor mounting quite stiff, it can be expected that the symptomatic resonance of the sensor mounting will be very high. Thus, at first sight it appears that a properly designed sensor should be capable of self-diagnosis of such a fault by itself.

However, we should also consider the way in which faults in the structure are likely to manifest themselves. Imagine that the beam shown above suffers an incomplete fracture. This results in two beams, connected by a compliance, giving an electrical analogue shown in Figure 3. This is similar to that for the loosely mounted sensor, with the exception of the coupling compliance. Now there are two beam resonances of concern, but in most cases it is likely that they will be at a much lower frequency than the symptomatic resonance of the sensor. The only case in which there might be some scope for misidentification is if the sensor is located close to the end of the beam and the fracture is located close to the sensor, resulting in a very high resonant frequency for the piece of the beam close to the sensor. If this frequency is in the range of that expected from a loose sensor, then the difference between the two signatures is the compliance linking the two resonant systems which occurs if the fault is in the beam, rather than the sensor mounting. The effect of this compliance is shown in the frequency plot in Figure 3. It is possible that a suitably designed diagnostic could detect such a characteristic, but the presence of information from adjacent sensors as to the magnitude of such a resonance would certainly make diagnosis easier.

Case 2. An extreme case of the above is if the sensor detaches completely from the structure under test, in which case no resonances will be detected. Such a condition may occur simply because there is no stimulus to cause resonance. Information from other sensors in the locality can be used to identify whether this is the case, a detached sensor being indicated by a single sensor detecting no resonances surrounded by ones which do. Complete loss of function of a sensor, for instance due to loss of power supply, would cause similar symptoms, except that there would not even be residual noise detected.

Case 3. Another related case is if the sensor housing suffers some structural damage. Such a problem may include a number of cases, the simplest of which is a small part of the housing becoming partially detached. Once again this takes the form of a small mass attached to the sensor by a compliance. The difference from the case discussed earlier is that the mass is still smaller and so the resonance will be higher, thus the discussion above applies, although the two symptoms should be more easily separated.

Case 4. Changes in ambient conditions can be detected at the sensor, if properly equipped and compensated for appropriately.

Case 5. Detection of parameter drift of an individual sensor, or of internal damage to its structure would cause would depend on detection of variation of the output of that sensor from its neighbors, once again in a way which could not be confused with symptoms of the structure under test. The nature of these variations would need to be characterized to be detected.

The faulty and healthy sensor signatures can be analyzed with a view to extracting their characteristic features. An example of such analysis is proposed below.

4 NEURAL NETWORK BASED SENSOR HEALTH DIAGNOSIS

Following on from the discussion in section 3, it becomes clear that, the sensor self-validation on the basis of its output only is an impossible task. A set of rules based on physical principles can be deduced for the data expected from neighboring sensors. (For example, acceleration measurements from sensors situated at adjacent locations along a cantilever will only be permitted to be different within certain limits imposed by the expected accelerations to which the object is subjected.) The application-related reasoning should be kept to a minimum, in order to maintain the overall generality of the method.

For the study presented here, it is assumed that we have two neighboring sensors, S1 and S2, measuring, for simplicity, the same acceleration. Note that the fact that both sensors sense the same acceleration does not bring any limitations to the methodology under development; the important aspect is that the output of two sensors (at least, as it is going to be discussed later) are needed to accomplish the task. The diagnosis system under design is associated with sensor S1 (DIAGNOSTIC NETWORK 1) and the contextual information is provided by sensor S2 (Figure 4). The input acceleration for the sensors ($a1(k)=a2(k)$) is a filtered white noise signal, with a frequency range of 0-100Hz, sampled with a sampling rate of 4kHz. One sensor fault only was considered at this stage, corresponding to Case1 in the previous section. The scaled input acceleration, healthy sensor output and faulty sensor output are shown in Figures 5a and 5b. The diagnosis module to be designed will consist of an ANN, whose task is to identify the healthy/faulty response of sensor S1 in the following situations: S1, S2 – healthy; S1 healthy, S2 faulty; S1 faulty, S2 healthy.

The ANN proposed is of a tap-delayed feed-forward type, with two hidden layers, trained by dynamic backpropagation (with a momentum term and a variable learning rate). Use is made of tap-delayed-lines (TD) in order to incorporate the dynamic behavior of the sensor into the model. Two delay units are necessary, to generate the one-step and two-steps back sensor output signals (*S1(k-1), S1(k-2), S2(k-1), S2(k-2)*), respectively. Hence, the present and the past values of the sensors S1 and S2 outputs form the input vector to the neural network. The network output represents the Healthy/Faulty condition of S1 at any instant of time (*k*). Since no feedback loop exists in the model, static error backpropagation (BKP) can be used to adjust the network parameters.

Based on these considerations, the electrical equivalent circuits in Figure 1 and 2 were simulated in SPICE in order to gather the input-output network training data. The ANN was trained and tested using Matlab. The best network performance (in terms of least false alarms and highest correct diagnosis rate) was obtained with a 6x31x17x1 network architecture. The network performance on a test set (produced under the same conditions as the training set) is shown in Figure 6. The continuous line represents the correct diagnosis expected from the sensor, with +0.99 identifying the Healthy condition and –0.99 identifying the Faulty condition; the dotted line represents the actual DIAGNOSTIC NETWORK 1 response. A "zero level" decision boundary would mean that the network correctly assesses the sensor's health, with two exception, both corresponding to the case where the sensor S1 is Faulty and S2 is Healthy.

It has therefore been possible to design a working self-diagnosis module for the acceleration sensor considered, on the basis of its own output signal and a minimum of contextual information, non-application based, provided by one neighboring sensor. The results obtained encourage the continuation of this line of research towards multiple faults diagnosis. It should be noted here that a single neighbor as context might not be sufficient for such a task.

5 CONCLUSIONS

The paper discussed the suitability and feasibility of enhancing the reliability of microsensors by adding an on-chip self-diagnosis capability. The approach taken is based on Artificial Intelligence techniques and sensors with no accessible internal signals are taken as an example. Some common acceleration sensor faults are considered and an indication is given of the manner in which these faults can be detected and isolated, either on an individual sensor basis or based on cooperative work within a sensor network. The design requirements for such self-diagnosable measurement systems are set and initial self-diagnosis implementation issues are tackled. A Self-diagnosis sensor module was designed which is able to successfully assess its state of health in respect to one fault condition. Work is in progress to enhance the module's ability to detect several other fault conditions.

REFERENCES

[1] Gaura, E. Neural Network Techniques for the Control and Identification of Acceleration Sensors. PhD Thesis, Coventry Univ., UK, 2000

[2] NASA TechTracS. System Fault Detection and Accommodation, Technology Opportunity Sheet, http://technology.larc.nasa.gov/scripts/nls_ax.dll/twDispTOPSItem(111;TOP3-2;0;1), 2000

[3] Patton, R., Uppal, F., Lopez-Toribio, C. Soft computing approaches to fault diagnosis for dynamic systems: a survey, IFAC Safeprocess, Hungary, Proceedings, pp. 298-311, 2000.

[4] Henry, M.P. Validating data from smart sensors, J. of Control Engineering, vo.41, no.9, pp. 63-66, 1994

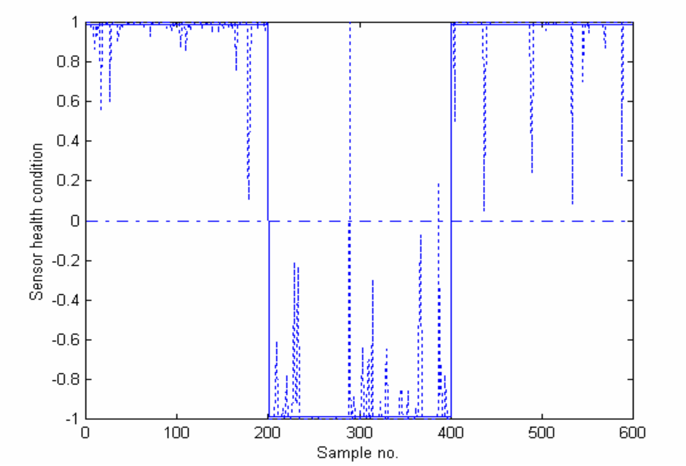
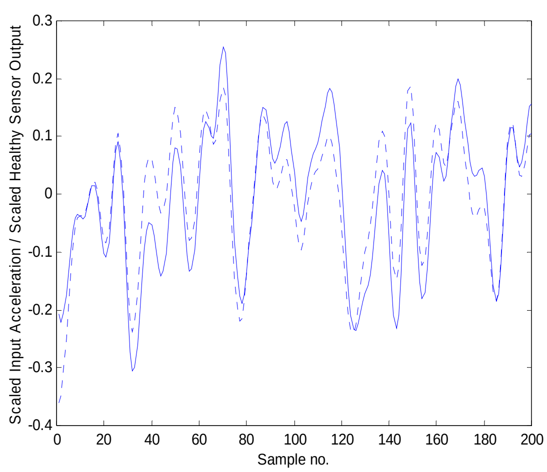


Figure 6: Diagnostic network performance



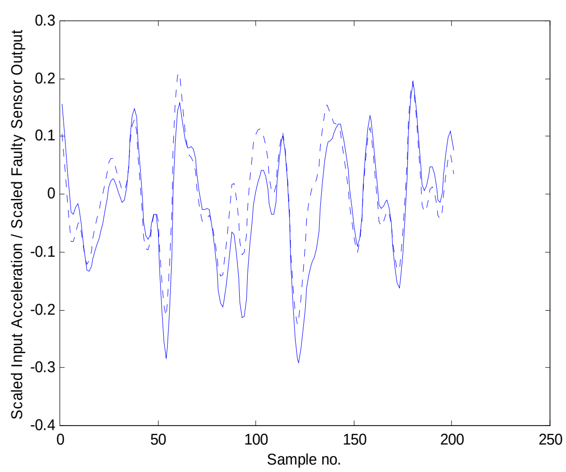


Figure 5a: Scaled Input Acceleration; Scaled Healthy Sensor Output

Figure 5b: Scaled Input Acceleration; Scaled Faulty Sensor Output

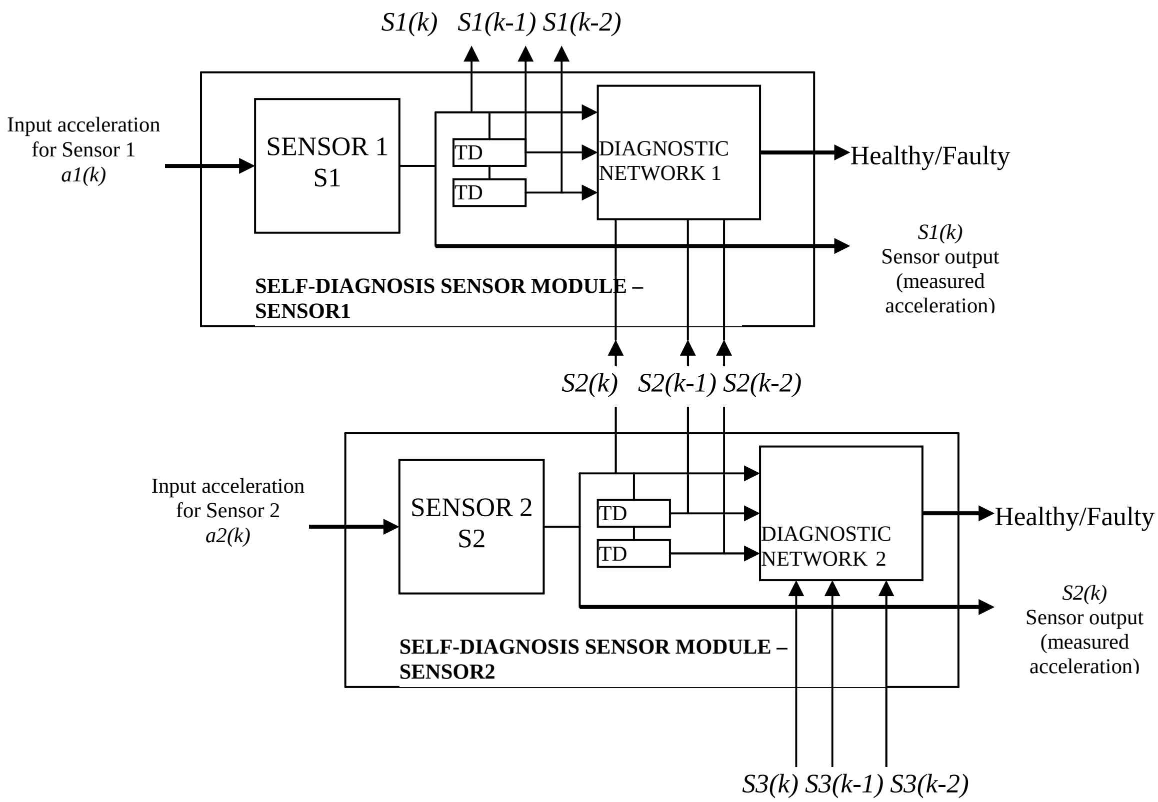


Figure 4: Block diagrams of two neighboring self-diagnosis sensors (*a1(k) = a2(k),* in this study)