Sensor Set Decomposition for Enhanced Distributed Sensor Fault Isolability of Marine Propulsion Systems

Kougiatsos, Nikos; Reppa, Vasso

DOI
10.1016/j.ifacol.2024.07.192

Publication date
2024

Document Version
Final published version

Published in
IFAC-PapersOnline

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-4 (2024) 49–54

Sensor Set Decomposition for Enhanced Distributed Sensor Fault Isolability of Marine Propulsion Systems ⋆

Nikos Kougiatsos * **Vasso Reppa** *

* *Department of Maritime and Transport Technology, Delft University of Technology, 2628CD, the Netherlands, (e-mail: {n.kougiatsos, v.reppa} @ tudelft.nl).*

Abstract: This paper proposes a greedy stochastic optimization algorithm for the sensor set decomposition used in the sensor fault monitoring of marine propulsion systems, based on fault isolability criteria. These criteria are expressed mathematically in terms of the number of unique columns in the theoretical fault signature matrices (FSMs) used during the sensor fault isolation process. Due to the large scale and complexity of marine propulsion plants, the diagnostic layer follows a distributed architecture with a combinatorial logic used for fault isolation in two cyber levels; the local and global decision logic. As a result, the FSMs of both levels are formulated as an integrated optimization problem. Each solution regarding the sensor set decomposition is then used to generate the respective distributed monitoring architecture, using semantic (qualitative) knowledge for the propulsion plant. Thus, the need for an analytical model of the plant is removed. Moreover, based on the design of the distributed monitoring architecture, the respective theoretical FSMs (quantitative) are automatically generated and used for the evaluation of the objective function. Finally, simulation results are used to illustrate the application of the greedy stochastic optimization algorithm and its efficiency.

Keywords: semantic knowledge models, FDI by means of structural properties, distributed fault diagnosis, sensor faults, complex systems, Artificial Intelligence methods, safety of marine systems

1. INTRODUCTION

Safety is a prerequisite for the design of modern marine vessels and a basic pillar for the development of future autonomous ships [de Vos et al. (2021)]. A vessel can be considered as a system of systems, each of which contains many subsystems that support its operation. Marine systems are typically designed with specific fuel, actuation and sensing capabilities. Their components are sourced by different manufacturers, with various specifications and in most cases their settings and configuration parameters are under tight-lock. The most complex and safety-critical vessel subsystem is the propulsion system.

To ensure onboard safety, novel and efficient diagnostic systems must be developed for marine vessels. These systems usually rely on the existing sensor information to make decisions. Since all onboard hardware sensor devices are sourced from different manufacturers, a collaborative framework between hardware sensors is required for diagnosis purposes. However, any given hardware sensor can suffer from faults, an issue that has not been properly addressed in maritime literature [Kougiatsos and Reppa (2024)]. Most papers in literature [Wu et al. (2006)] explore the use of hardware redundancy to improve the efficiency of the monitoring system and recover from sensor faults. However, this strategy leads to greater hardware installation and maintenance costs. Instead, analytical redundancy tools like virtual sensors could be considered [Kougiatsos and Reppa (2022)]. In both cases, the optimality of the diagnosis approach's design should be examined.

The optimisation of the fault diagnosis process is often associated with the optimal choice of the sensor set decomposition. The sensor set decomposition problem aims to determine the optimal number and composition of sensor subsets, stemming from the starting sensor set, in order to enhance the isolation of multiple sensor faults [Reppa et al. (2016)]. This is especially necessary in large networks of cyber-physical interconnected systems, as in the case of marine propulsion systems, where the isolation of multiple sensor faults is really difficult or even infeasible with a single monitoring module. The objective function in these problems is based on nonlinear observer stability and fault isolability objectives rather than cost. In addition, applications on centralised monitoring architectures are mostly discussed in literature [Reppa et al. (2016)], which would however result in high computational burden for large-scale applications such as marine vessels.

In marine literature, the sensor set decomposition problem has not received much attention. Stoumpos and Theotokatos (2022) propose a Unified Digital System for diagnosis and health management of dual-fuel engines.

⋆ This publication is part of the project READINESS with project number TWM.BL.019.002 of the research programme "Topsector Water & Maritime: the Blue route" which is partly financed by the Dutch Research Council (NWO).

10.1016/j.ifacol.2024.07.192

Due to the inherent complexity of marine engine modelling, the digital twin was realised using the tools, libraries and functionalities of a commercial program. The sensors were grouped in subsets according to the system decomposition (e.g., Gas Valve Unit, Exhaust waste gate) and the type of faults considered (e.g. sensor, actuator faults). However, the optimality of this sensor set decomposition was not discussed. Other papers [Wohlthan et al. (2021)], deal with modelling complexity by using part of the complete fuel engine model for diagnosis purposes and only a limited selection of sensors. Due to the limited size of the sensor set, the sensor set decomposition problem has not been addressed either.

In previous work [Kougiatsos and Reppa (2024)], the authors proposed a distributed sensor fault diagnosis framework for marine internal combustion engines and the diagnosis scheme was assessed using both detectability and isolability performance metrics. However,the sensor set decomposition in the distributed monitoring architecture was not optimal and was related to the physical system decomposition. In order to deal with the inherent modelling complexity in marine propulsion systems, the design of a semantic database of system and automation components was also previously presented in [Kougiatsos et al. (2023)].

The objective of this paper is the design of an optimisation algorithm for the sensor set decomposition to enhance the fault isolability of sensor faults in marine propulsion systems. The optimization criteria are mathematically described by the number of unique columns in the theoretical Fault Signature Matrices (FSMs) used for the isolation of multiple sensor faults. This paper considers a marine hybrid propulsion system, composed of a combination of a marine internal combustion engine (mechanical power) and an induction motor (electrical power). Due to the large scale and complexity of marine propulsion installations, a distributed monitoring architecture is considered, (see Section 2). The exchange of information between the destributed monitorng agents mimics the interconnection dynamics of the real system. For this reason, fault isolation occurs in two levels; the local and global decision logic. To automatically generate the various FSMs and interconnections between the monitoring agents in the distributed monitoring architecture, a qualitative modelling technique is employed based on semantic knowledge and is presented in Section 3. The semantic database includes information on the various considered hardware automation components (e.g., hardware sensors, controllers) and a knowledge graph to visualise their connections. Analytical redundancy considerations are also included in the semantic description in the form of virtual sensors [Kougiatsos and Reppa (2022)], used to construct more Analytical Redundancy Relations (ARRs) and improve the isolability of multiple faults. The integrated optimization problem considering both isolation levels is then formulated in Section 4 and a greedy stochastic optimisation algorithm is proposed for its solution. The greedy optimiser is applied in a marine propulsion case study in Section 5 while some concluding remarks are provided in Section 6.

The main contribution of this research work is the optimisation of the sensor set decomposition for enhanced isolability in multiple sensor fault scenarios affecting marine propulsion systems. Compared to the fault diagnosis

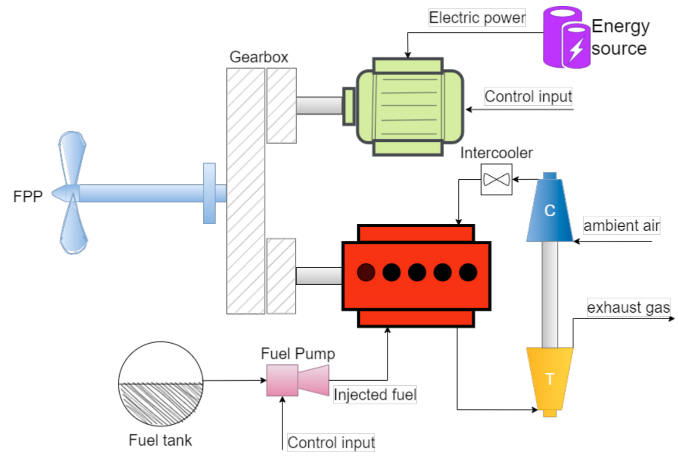

Fig. 1. Marine hybrid propulsion system layout. The system involves greatly heterogeneous dynamics and a considerable number of interconnections, adding to its complexity.

literature, this paper optimises the grouping of sensors in a distributed scheme. In order to address the complexity in modelling the marine propulsion system while still providing quantitative inputs for the optimiser (FSMs), the use of a semantic database of components (qualitative information) is proposed instead of analytical models. Finally, the semantic database is enriched with the use of virtual aside from the hardware sensors. The analytical redundancy consideration in turn limits the required hardware as well as its installation and maintenance costs.

2. DISTRIBUTED SENSOR FAULT DIAGNOSIS

In Fig. 1, an example layout of a marine hybrid propulsion system is shown. Due to the large scale and complexity of marine propulsion installations, the application of a distributed monitoring architecture has already been proposed in [Kougiatsos and Reppa (2024)]. In this setup, monitoring agents $\mathcal{M}^{(I)}$, $I = 1, \cdots, N$ are designed, each consisting of N_I modules $\mathcal{M}^{(I,q)}$, $q = 1, \cdots, N_I$. Every agent $\mathcal{M}^{(I)}$ monitors a set of sensors $\mathcal{S}^{(I)}$, which is a subset of the global set of sensors denoted by $\mathcal{S}$. This section will introduce the basics of the diagnosis process.

Distributed detection Every monitoring module $\mathcal{M}^{(I,q)}$ obtains a decision based on one or more Analytical Redundancy Relations (ARRs). The j-th ARR can be defined as:

$$\mathcal{E}_j^{(I,q)} : \epsilon_{yj}^{(I,q)}(t) \in \mathbb{E}^{(I,q)}, \tag{1}$$

where $\epsilon_{yj}^{(I,q)}$ signifies the residual and $\mathbb{E}^{(I,q)}$ denotes the bounds of this residual, often noted as the thresholds. The set of ARRs used by the module $\mathcal{M}^{(I,q)}$ are defined as: $\mathcal{E}^{(I,q)} = \cup_{j\in\mathcal{J}^{(I,q)}}\mathcal{E}_j^{(I,q)}$. where $\mathcal{J}^{(I,q)}$ is an index set, defined as $\mathcal{J}^{(I,q)} = \{j : S^{(I)}\{j\} \in S^{(I,q)}\}$.

Each module's decision on the occurrence of sensor faults, in the case of permanent faults is defined as:

$$D^{(I,q)}(t) = \begin{cases} 0, & \mathcal{E}^{(I,q)} \text{ is valid} \\ 1, & otherwise \end{cases} \tag{2}$$

Sensor fault isolation Sensor fault isolation relies on a two-step combinatorial decision logic; the local and the global decision logic. In the local decision level, the decisions $D^{(I,q)}$ are collected into a binary decision vector $D^{(I)} = [D^{(I,1)}, D^{(I,2)}, \cdots, D^{(I,N_I)}]$ representing the local decision of the monitoring agent $\mathcal{M}^{(I)}$. This vector is compared to the columns of a binary fault signature matrix (FSM) $F^{(I)}$ consisting of N_I rows and $N_{C_I} + 2$ columns where $N_{C_I} = 2^{|S^{(I)}|} - 1$, $|S^{(I)}|$ denotes the cardinality of the sensor subset $S^{(I)}$. The local decision vector $D^{(I)}$ corresponding to the monitoring agent $\mathcal{M}^{(I)}$ is consistent with the i-th column of the binary fault signature matrix $F^{(I)}$, denoted as $F^{(I)}_{qi}$, $q = 1, \cdots, N_I$ if and only if $D^{(I,q)} = F^{(I)}_q$, $\forall q = 1, \cdots, N_I$. The local diagnosis set of the agent is then defined as:

$$\mathcal{D}^{(I)}_s = \{\mathcal{F}^{(I)}_{ci} : F^{(I)}_i = D^{(I)}(t)\}, \quad (3)$$

where $\mathcal{F}^{(I)}_{ci}$ is a consistency test between the i-th column of the fault signature matrix $F^{(I)}$ denoted as $F^{(I)}_i$ and the local agent decision vector $D^{(I)}$.

In order to isolate the propagation of the sensor fault effects, the global agent collects the decisions of the monitoring agents $\mathcal{M}^{(I)}$, denoted as:

$$D^{(I)}_\chi(t) = \begin{cases} 0, & f^{(I)}_\chi \notin \mathcal{D}^{(I)}_s(t) \textit{ and } f^{(I)}_p \notin \mathcal{D}^{(I)}_s(t) \\ 1, & \textit{otherwise}, \end{cases} \quad (4)$$

where $f^{(I)}_p$ accounts for the sensor faults propagated from the agent to the neighboring agents and $f^{(I)}_\chi$ corresponds to the sensor faults propagated from the neighboring agents to the local agent.

A global agent $\mathcal{G}$ collects the decisions on the propagation of sensor faults from the N local agents in a global decision vector $D_\chi(t) = [D^{(1)}_\chi(t), D^{(2)}_\chi(t), \cdots, D^{(N)}_\chi(t)]$. A comparison is then made with the columns of a global binary sensor FSM F^χ consisting of N rows and $N_C = 2^p - 1$ columns where $p \leq \sum_{I=1}^{N}\{p_I\}$, p_I is the length of $f^{(I)}_\chi$.

The local multiple sensor fault isolability becomes complicated as the number of sensors in $\mathcal{S}^{(I)}$ increases. To handle this complexity, the objective of this paper is to design an optimisation algorithm for the decomposition of the sensor set $\mathcal{S}$ in $\mathcal{S}^{(I)}$, $I = 1, \cdots, N$ subsets and the automated generation of the theoretical FSMs $F^{(I)}$, F^χ and agents' architecture ($\mathcal{M}^{(I)}$, $\mathcal{G}$) used in the distributed monitoring of marine propulsion systems, based on fault isolability criteria. In this paper we consider that each monitoring module $\mathcal{M}^{(I,q)}$, $I = 1, \cdots, N$ only uses one ARR, meaning that $q = 1, \cdots, |\mathcal{S}^{(I)}|$.

3. SEMANTIC MODELLING OF MARINE PROPULSION SYSTEMS

In this work we propose the use of a qualitative method to handle the complexity of the modelling of the hybrid propulsion system shown in Fig. 1. In previous work [Kougiatsos et al. (2023)], a semantic database of vessel components was introduced. The semantic database includes the components database where the semantic information about the system components is stored and the knowledge graph, a tool that helps visualize the connections between the different hardware and cyber components.

3.1 Components database ($\mathcal{F}$)

All the physical (plant, hardware sensors, actuators) and cyber (controllers, virtual sensors, monitoring agents) components can be semantically described by means of their type (e.g. sensor), their input (e.g. injected fuel) and output (e.g. engine torque) as well as their units. New components can be then easily appended in the database through the operation of semantic annotation [Kougiatsos et al. (2023)] while components can be also removed during the vessel's lifecycle. This process is done through the actions of the system designers and/or involved manufacturers. The component database $\mathcal{F}$ used in this paper can be analyzed as follows:

$$\mathcal{F} = \mathcal{F}_p \cup \mathcal{F}_a \cup \mathcal{F}_c \cup \mathcal{F}_s \cup \mathcal{F}_e \cup \mathcal{F}_y \cup \mathcal{F}_u \cup \mathcal{F}_v, \quad (5)$$

where $\mathcal{F}_a$, $\mathcal{F}_c$, $\mathcal{F}_s$, $\mathcal{F}_e$, $\mathcal{F}_y$, $\mathcal{F}_u$ denote the set of "actuators", "controllers", "sensors", "state-estimators", "pre-control functions" and "post-control functions" respectively, $\mathcal{F}_p$ denotes the physical plant components set (e.g. compressor, turbine, gearbox etc.).

Virtual sensors are in general multi-input multi-output cyber entities, which can be used instead of extra hardware sensors. The set of virtual sensors in the semantic database $\mathcal{F}$ is denoted as $\mathcal{F}_v$. Their design is usually based on analytical model information of the plant. In particular, three types of virtual sensors have been derived for Differential-Algebraic systems in [Kougiatsos and Reppa (2022)]; dynamic, static and Set Inversion via Interval Analysis (SIVIA)-based virtual sensors. The first type uses measurements from its hardware sensor counterpart alongside other hardware sensors, while the other two types do not.

3.2 Knowledge Graph (G)

The knowledge graph of the plant is a graph of the system formed using the available knowledge about its operation (semantic description of vessel components). In [Kougiatsos et al. (2023)], an automated knowledge graph algorithm has been proposed. This results in the knowledge graph of the plant hereby denoted as $G = \{V, E, Y\}$ where V denotes the vertices of the graph (database components), E denotes the edges between the vertices of the database (connections constructed by the knowledge graph tool using the available semantic information) and Y denotes the qualitative information carried by each edge (e.g. 'fuel' is used as a connection between the fuel tank and the fuel pump). The system interconnections included in the knowledge graph can be used to automatically configure the cyber connections between the monitoring agents in the distributed monitoring architecture.

4. DIAGNOSTIC SYSTEM DESIGNER MODULE

As previously discussed in Section 2, the design of the diagnosis cyberlayer depends on the decomposition of sensors in subsets $\mathcal{S}^{(I)}$, $I = 1, \cdots, N$ and their assignment to monitoring agents $\mathcal{M}^{(I)}$. In this work, the physical system and its associated sensor set will be decomposed based on the maximisation of the ability to isolate sensor faults. This property can be expressed using the local $F^{(I)}$ and global F^χ theoretical fault signature matrices. The challenging part of the distributed architecture is the decision

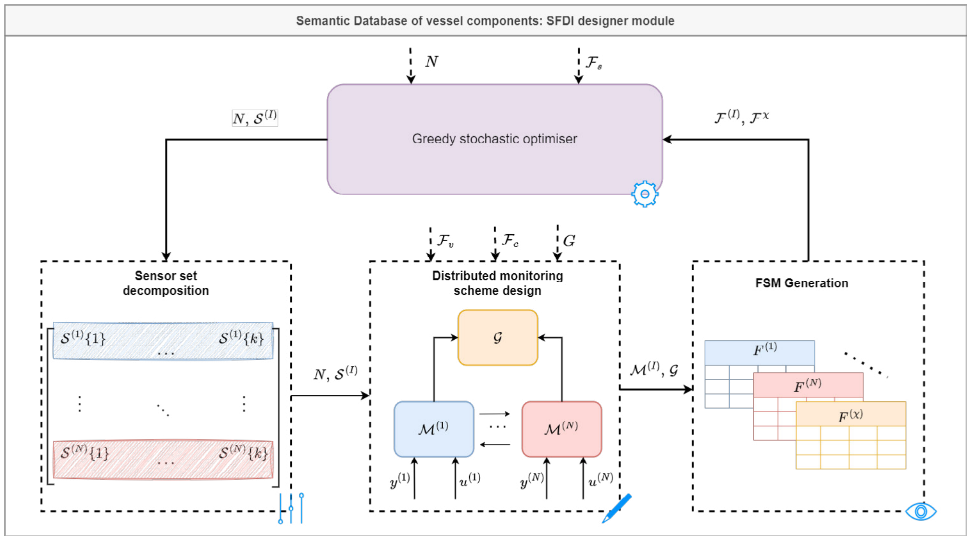

Fig. 2. The greedy stochastic optimiser determines the composition of sensor sets $\mathcal{S}^{(I)}$, $I = 1, \cdots, N$ that maximises the isolability targets, as those can be expressed using the theoretical fault signature matrices ($F^{(I)}$, F^{χ}). The inputs to the greedy stochastic algorithm in each different step are shown with dashed lines. The process described in this figure is implemented as a functionality of the semantic database. As such, the semantic information for hardware ($\mathcal{F}_s$), virtual sensors ($\mathcal{F}_v$) and controllers ($\mathcal{F}_c$) can be used. In addition, the knowledge graph G is available.

complexity introduced by the combinatorial decision logic. The fault signature of two faults might be the same in one FSM but different in another one, so these two faults can be isolated from one another. Moreover, sensor faults can be propagated between different local agents through their interconnections and multiple decisions are needed to exclude the possibility of certain fault combinations occurring. Thus, a different formulation of the sensor set decomposition problem is required to handle the challenges of the distributed monitoring architecture and will be presented next. The diagnostic system designer module presented in this Section is implemented as an additional utility of the semantic database. As a result, this tool can benefit from the semantic information (see Section 3) regarding the hardware sensors ($\mathcal{F}_s$), the virtual sensors ($\mathcal{F}_v$), the controllers ($\mathcal{F}_c$) and the knowledge graph (G) in order to automatically generate sensor set decompositions and the respective distributed monitoring architectures. The process of decomposing the original sensor set into subsets occurs offline.

4.1 Optimisation of sensor monitoring decomposition

The optimisation problem for the sensor fault diagnosis process, in a distributed monitoring architecture, can be expressed as follows:

$$\max_{N,S^{(I)}} \quad \rho\left(\Phi(F^{(I)})\right) + \rho\left(F^{(\chi)}\right) \tag{6}$$

$$\text{s.t.} \quad S^{(I)} \cap S^{(J)} = \emptyset \ \ \forall I \neq J, \tag{7}$$

$$\cup_{I=1}^{N} \mathcal{S}^{(I)} \subseteq S, \tag{8}$$

$$N \leq N_{max} \tag{9}$$

$$\mathcal{S}_R \subseteq \cup_{I=1}^{N} \mathcal{S}^{(I)}, \tag{10}$$

where the notation $\rho(A)$ signifies the rank of matrix A and $\Phi : F^{(I)} \mapsto F^c$ is a mapping transforming the local fault signature matrices to one equivalent sensor fault signature matrix with the total number of rows and the total number of columns of all local matrices $F^{(I)}$. **The objective function** (6) aims to mathematically express the sensor fault isolability property of the (under design) distributed monitoring architecture in terms of the number of unique columns (or rank) of the matrices $F^{(I)}$, F^{χ}. **Constraint** (7) signifies that each sensor can be assigned to only one monitoring agent, though its measurement may be transmitted between other agents as well. **Constraint** (8) is used so that sensors may not be used if they make no difference in the diagnosis process but they are selected from a limited pool of available hardware sensors. To implement the above constraints, the optimizer uses the available semantic information for hardware sensors ($\mathcal{F}_s$) as shown in Figure 2.

Moreover, **constraint** (9) aims to limit the size of the design space by keeping the number of created agents N bounded by a parameter N_{max}. In order to construct residuals in the graphs, the hardware sensor vertices ($\mathcal{F}^s$) belonging to the selected $\mathcal{S}^{(I)}$, $I = 1, \cdots, N$ need to be combined with similar "virtual sensor" vertices ($\mathcal{F}^v$) (e.g. if the shaft speed sensor is chosen, the virtual sensor for shaft speed needs to be coupled). Since virtual sensors require certain inputs from hardware sensors ($\mathcal{F}_s$) and controllers ($\mathcal{F}_c$) to be functional, **constraint** (10) aims to impose the selection of these hardware sensors in the designed division of sensors. The set of virtual sensor requirements $\mathcal{S}_R \subseteq \mathcal{F}_s$ can be defined as the semantic inputs of the virtual sensors and is also used to determine the interconnections between the monitoring agents $\mathcal{M}^{(I)}$ based on the knowledge graph G. Thus, the distributed monitoring architecture is automatically constructed, as shown in Figure 2. Finally, using the information about the interconnections between the agents $\mathcal{M}^{(I)}$, $\mathcal{G}$ and considering that each monitoring module $\mathcal{M}^{(I,q)}$ only employs one ARR (i.e. the number of modules of each agent $I \in [1, N]$ are $q = 1, \cdots, |\mathcal{S}^{(I)}|$), the local and global FSMs $F^{(I)}$, F^{χ} are also automatically generated.

4.2 Greedy stochastic optimisation algorithm

A heuristic search algorithm that can be used to solve the optimisation problem formulated in (6)-(10) is a greedy algorithm. This type of heuristic has been already used with success for a similar problem in [Jung et al. (2020)], proving a suitable candidate for our problem. However, the main issue associated with greedy algorithms is that they are deterministic and the solution greatly depends on the initial conditions of the search. As a result, the solution provided by the algorithm may be far from the global optimum. In order to mitigate this risk, in this research work, a greedy stochastic algorithm is used instead. The algorithm tunes the number of sensor subsets $\mathcal{S}^{(I)}$ ($\subseteq \mathcal{S} \subset \mathcal{F}_s$) and their hardware sensor composition, thus generating different distributed monitoring architectures. The local and the global FSMs, $F^{(I)}$ and F^{χ}, are then derived and used for the calculation of the objective function in (6). The algorithm is repeated until convergence, as indicated in Fig. 2. In order to represent the search space of all possible sensor divisions in a structured manner, a lattice representation is used. In every run of the algorithm, each monitoring agent is first assigned randomly a branch of the lattice tree, satisfying the constraints (7)-(10), similar to [Jung et al. (2020)]. The monitoring agents then sequentially make decisions on whether to drop, add, exchange a hardware sensor with another agent or do nothing. The objective of each agent at its decision step is the maximisation of the objective function in (6), considering the decisions of the previous agents and supposing that the following agents will opt to maintain their sensor sets as is [Konda et al. (2022)]. In order to assess the gains of the different options, 5 random removals, additions and exchanges of sensors are considered. The execution of the algorithm stops when all monitoring agents opt to maintain their sensor set division. The main improvements of our algorithm compared to the relevant literature [Jung et al. (2020)] are **(a)** its suitability for highly complex systems by combining quantitative and qualitative tools; and **(b)** the consideration of distributed monitoring architectures and associated challenges.

5. SIMULATION RESULTS

In this section we apply the diagnostic system designer module to a hybrid marine propulsion system, such as the one shown in Figure 1. In total, 12 hardware sensors and 13 virtual sensors are available. For brevity purposes the sensors will be referred to using their IDs in the Semantic Database, shown in Table 1. In order to assess the optimal number of monitoring agents, the greedy stochastic algorithm is executed for $N \in \{1, 2, 3, 4, 5\}$. The optimal sensor set division per each number of agents N is defined as the one with the maximum cost value defined in (6) after running the algorithm 30 times.

Table 1. Sensor IDs as vertices in the semantic database

ID	Sensor	ID	Sensor
18	Fuel injection	24	Turbine temperature
19	Cylinder pressure	25	Compressor pressure
20	Cylinder temperature	26	Compressor temperature
21	Engine torque	27	Intercooler temperature
22	Exh. manifold pressure	28	Shaft speed
23	Exh. manifold temperature	29	Motor torque

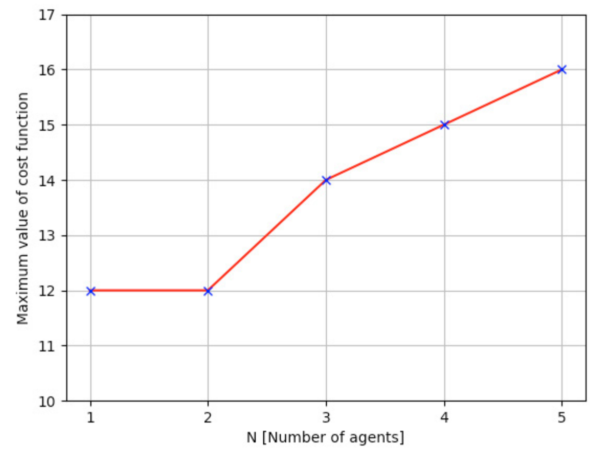

Fig. 3. Depiction of optimal solution costs (shown in the y axis) per each value of N. The data points are shown with a blue x marker.

The results of the greedy stochastic optimisation algorithm are shown in Figure 3. In each run, the algorithm starts from a random sensor division with length equal to the number of monitoring agents. The blue points in Figure 3 then correspond to the maximum value of cost, meaning the maximum number of unique columns in the FSMs achieved by the integrated optimisation problem combining both isolation levels in (6), obtained for all the runs in a specified value of N. As observed, the centralised ($N = 1$) and distributed with two monitoring agents ($N = 2$) configurations result in the same maximum cost value. This is due to the high inter-connectivity of the physical system (marine hybrid propulsion system), which in the case of the two-agent configuration results in the monitoring agents being fully connected (the matrix F^{χ} has only one unique column). The cost value significantly increases when considering the $N = 3$ agent configuration while smaller increases are seen when $N = 4$ or $N = 5$. The maximum cost is encountered in the 5-agent distributed configuration. For this specific data point the optimal sensor divisions are the following: $\mathcal{S}^{(1)} = \{19, 20, 21, 23, 25, 27, 29\}$, $\mathcal{S}^{(2)} = \{18, 28\}$, $\mathcal{S}^{(3)} = \{24\}$, $\mathcal{S}^{(4)} = \{22\}$ and $\mathcal{S}^{(5)} = \{26\}$. The virtual sensors are realised in the monitoring agent that their hardware sensor counterpart has been assigned to. Suppose that we want to visualise the optimal sensor set decomposition for $N = 5$ by extrapolating it as a decomposition of the physical plant shown in Figure 1 in multiple systems; the result is shown in Figure 4. As observed from Figure 4, some systems are overlapping ($\Sigma^{(1)}$, $\Sigma^{(3)}$, $\Sigma^{(4)}$, $\Sigma^{(5)}$) due to their hardware sensors being assigned to multiple non-overlapping sensor sets. Finally, the resulting local FSM for the first monitoring agent $\mathcal{M}^{(1)}$ can be seen in Table 2 while the global FSM F^{χ} is given in Table 3. In each matrix, every different column represents the theoretical signature of a fault on the sensor ID given in the header of the column. For brevity purposes, only single fault columns are shown. By using virtual sensors, the sensitivity of the resulting analytical redundancy relations is the same both for local and propagated sensor faults. As a result, no ambiguity is taken into consideration in the FSMs, as opposed to [Kougiatsos and Reppa (2024)].

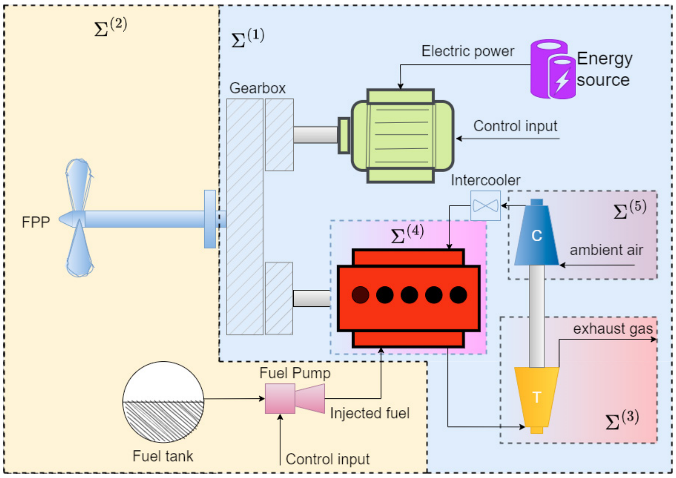

Fig. 4. Marine hybrid propulsion system physical decomposition based on the optimisation of the sensor fault diagnosis. The 5 subsystems are shown with different colors. The subsystems $\Sigma^{(3)}$, $\Sigma^{(4)}$, $\Sigma^{(5)}$ are overlapping with subsystem $\Sigma^{(1)}$. For each subsystem $\Sigma^{(I)}$, $I = 1, 2, \cdots, 5$ a monitoring agent $\mathcal{M}^{(I)}$ is created.

Table 2. Part of the Sensor Fault signature matrix $F^{(1)}$ of $\mathcal{M}^{(1)}$

	18	19	20	21	22	23	24	25	26	27	29
$\mathcal{E}^{(1,1)}$	0	1	0	0	0	0	0	0	0	0	0
$\mathcal{E}^{(1,2)}$	0	1	1	0	0	0	0	0	1	1	1
$\mathcal{E}^{(1,3)}$	0	1	0	1	0	0	0	0	1	1	1
$\mathcal{E}^{(1,4)}$	0	1	1	1	0	1	1	0	1	1	1
$\mathcal{E}^{(1,5)}$	0	0	0	0	0	1	1	1	0	0	0
$\mathcal{E}^{(1,6)}$	0	0	0	0	0	0	0	0	0	1	0
$\mathcal{E}^{(1,7)}$	1	0	0	0	1	0	0	0	0	0	1

Table 3. Part of the Global Fault signature matrix of F^{χ} for a 5 agent distributed configuration

	18	19	20	21	22	23	24	25	26	27	29
$\mathcal{M}^{(1)}$	0	1	0	0	0	1	1	0	1	1	1
$\mathcal{M}^{(2)}$	1	1	1	1	1	1	1	1	1	1	1
$\mathcal{M}^{(3)}$	0	1	1	1	0	1	1	0	1	1	1
$\mathcal{M}^{(4)}$	0	1	0	0	1	0	0	0	1	1	1
$\mathcal{M}^{(5)}$	1	1	0	0	0	1	1	1	1	1	1

Based on the above results, a larger number N of decompositions of the sensor set $\mathcal{S}$ in $\mathcal{S}^{(I)}$, $I = 1, \cdots, N$ with an equally large amount of monitoring agents $\mathcal{M}^{(I)}$ seems to result in more isolable columns in the FSMs, at the cost of greater communication needed between the monitoring agents. Based on the automatically generated FSMs shown in Tables 2 and 3, we can see a large number of unique fault signatures. In particular, **9 out 11** unique columns are observed in Table 2 and **5 out of 11** in Table 3 considering the single faults case. Nonetheless, based on the global FSM in Table 3, the monitoring agents are very interconnected with **4 out of 11** faults affecting all 5 monitoring agents, **6 out of 11** faults affecting at least three agents and at all cases of faults **(11 out of 11)** affecting at least two monitoring agents. Moreover, although the sensor sets are not overlapping by design (due to constraint (7)), if we choose to decompose the system based on the sensor set division, the resulting systems might be overlapping, as shown in Figure 4.

6. CONCLUSION

In this paper, a distributed diagnostic system designer module was developed for the sensor set decomposition problem encountered in the distributed monitoring of marine hybrid propulsion architectures. Due to its inherent modelling complexity, the propulsion system was modelled using a qualitative approach, based on semantic information about its components and a knowledge graph to visualise their interconnections. The problem was then expressed in sensor fault isolability terms and a greedy stochastic optimiser was proposed for its solution. The obtained results from the case study indicated the efficiency of the algorithm, provided valuable insights on the optimal sensor set decomposition and highlighted the feature of automatically constructing the binary FSMs (quantitative) using the semantics (qualitative) modelling method.

REFERENCES

de Vos, J., Hekkenberg, R.G., and Valdez Banda, O.A. (2021). The Impact of Autonomous Ships on Safety at Sea – A Statistical Analysis. *Reliability Engineering and System Safety*, 210, 107558.

Jung, D., Dong, Y., Frisk, E., Krysander, M., and Biswas, G. (2020). Sensor selection for fault diagnosis in uncertain systems. *International Journal of Control*, 7179, 629–639.

Konda, R., Grimsman, D., and Marden, J.R. (2022). Execution Order Matters in Greedy Algorithms with Limited Information. *Proceedings of the American Control Conference*, 2022-June, 1305–1310.

Kougiatsos, N. and Reppa, V. (2022). A distributed virtual sensor scheme for marine fuel engines. *IFAC-PapersOnLine*, 55(31), 333–338.

Kougiatsos, N. and Reppa, V. (2024). A distributed cyber-physical framework for sensor fault diagnosis of marine internal combustion engines. *IEEE Transactions on Control Systems Technology*, 1–12. doi: 10.1109/TCST.2024.3378992.

Kougiatsos, N., Zwaginga, J., Pruyn, J., and Reppa, V. (2023). Semantically enhanced system and automation design of complex marine vessels. In *2023 IEEE Symposium Series on Computational Intelligence (SSCI)*, 512–518. IEEE.

Reppa, V., Polycarpou, M.M., and Panayiotou, C.G. (2016). Sensor fault diagnosis. *Foundations and Trends in Systems and Control*, 3(1-2), 1–248.

Stoumpos, S. and Theotokatos, G. (2022). A novel methodology for marine dual fuel engines sensors diagnostics and health management. *International Journal of Engine Research*, 23(6), 974–994.

Wohlthan, M., Schadler, D., Pirker, G., and Wimmer, A. (2021). A multi-stage geometric approach for sensor fault isolation on engine test beds. *Measurement: Journal of the International Measurement Confederation*, 168, 108313.

Wu, N.E., Thavamani, S., Zhang, Y.M., and Blanke, M. (2006). Sensor fault masking of a ship propulsion system. *Control Engineering Practice*, 14(11), 1337–1345.