



HAL
open science

Simulation-based fuzzy-rough nearest neighbour fault classification and prediction for aircraft maintenance

Yinling Liu, Tao Wang, Haiqing Zhang, Vincent Cheutet

► **To cite this version:**

Yinling Liu, Tao Wang, Haiqing Zhang, Vincent Cheutet. Simulation-based fuzzy-rough nearest neighbour fault classification and prediction for aircraft maintenance. *Journal of Simulation*, 2021, 15 (3), pp.202-216. <10.1080/17477778.2019.1680261>. <hal-02344426>

HAL Id: hal-02344426

<https://hal.science/hal-02344426v1>

Submitted on 31 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Simulation based fuzzy-rough nearest neighbour fault classification and prediction for aircraft maintenance

Yinling LIU^a, Tao WANG^b, Haiqing ZHANG^c and Vincent CHEUTET^a

^aUniversité de Lyon, INSA Lyon, Laboratoire DISP (EA4570), France;

^bUniversité de Lyon, Université Jean Monnet, Laboratoire DISP (EA4570), France;

^cSchool of Software Engineering, Chengdu University Information Technology, China.

ARTICLE HISTORY

Compiled January 31, 2020

ABSTRACT

Effective Fault Classification and Prediction (FCP) for aircraft maintenance helps improve the efficiency of maintenance. The traditional ways of the FCP mainly rely on the analysis of mechanical characteristics of components and the reliability-related data. Obtaining the fault-related data is usually prior to performing the analysis of fault diagnoses. However, acquiring the data especially the sensitive data in terms of security and confidentiality is extremely difficult. Therefore, in this paper, we propose an approach of combining Agent-Based Simulation System (ABSS) with the Fuzzy-Rough Nearest Neighbour (FRNN) approach, in order to implement the FCP for aircraft maintenance without the sufficient data. To do so, firstly, a framework for integrating the FRNN approach into ABSS is proposed. The concept and architecture models of FRNN-ABSS are used to describe FRNN-ABSS system. In the following, random and sequence strategies are designed for the FCP of the engine and an algorithm is developed to accomplish the integration of the FRNN approach and ABSS, aiming at automatically performing fault diagnoses. Finally, the experiments analysing the impact of different strategies on maintenance costs and service level have been conducted. The results show that the approach proposed achieved success for random and sequence strategies: 9.3% and 2.5% of maintenance costs have been saved; 4.17% and 12.5% of delayed flights have been changed into on-time flights.

KEYWORDS

Fault classification and prediction; Fuzzy-rough nearest neighbour; Agent-based system and simulation; Aircraft maintenance.

1. Introduction

The advanced technology and the complex system design have been broadly used in modern aircraft systems. Fault diagnoses become increasingly difficult because of the intractable faults and the complex connections between systems. Making use of an automatic classification scheme helps remove the need for the detailed analysis of the complex fault information (Mechefske and Mathew 1992). Early fault prediction is a proven technique in achieving high system reliability (Khoshgoftaar and Seliya 2003). Thus, effective FCP can significantly improve the efficiency of fault diagnoses, which finally helps achieve a highly efficient maintenance for aircraft. The function demands

of the FCP for aircraft maintenance include, but not limited to,

- building knowledge models on faults to store the fault information;
- automating the process of the FCP for providing a fast-response mechanism on aircraft maintenance under the big data environment;
- supporting the analysis of the FCP on the aircraft of new types, in order to prevent them from stopping working as far as possible.

The traditional rules of the FCP for aircraft maintenance are principally from air-worthy standards. These rules are refined from thousands of real cases of accidents about aircraft and the design theory(ASD 2019). Many new techniques have been attempted to improve the efficiency about the FCP for aircraft maintenance (Sahin et al. 2007; Jiao et al. 2017; Wang et al. 2018), whereas few of researchers have considered the case that the historic data for fault diagnoses is either not sufficient to be analysed or only valid for specific circumstances. This may raise big issues. For example, if the aircraft of a new type comes out, aircraft manufacturers simply have some limited data generated before delivering them to clients. They may suffer a lack of appropriate ways of fault diagnoses. Indeed, most of the rules still work, but how to investigate the functionalities and interactions of the new components remains unsolved. Besides, acquiring sensitive data like data for aircraft maintenance is exceptionally difficult. The lack of data is impeding researchers to study further the operational support of aircraft maintenance. Therefore, our study focuses on **how to improve the efficiency of the fault diagnosis for aircraft maintenance without sufficient data.**

The maintenance businesses of aircraft involve flight scheduling, maintenance strategies and planning, repairs, part supply and a number of stakeholders including OEM (Original Equipment Manufacturing), suppliers, airline, MRO (Maintenance, Repair and Overhaul), and airworthiness authority, etc. Stakeholders cooperate with each other in a distributed, synchronised and ephemeral way. For example, suppliers purchase parts independently while they are out of stock. At the same time, if necessary parts and auxiliary equipment for repairing some faulty components are ready, MRO workers will be scheduled. Maintenance managers will define strategies as soon as one maintenance is requested. Moreover, the maintenance of aircraft is mandated and monitored by regulatory authorities like FAA¹ (Federal Aviation Administration), CAA² (Civil Aeronautics Authority), etc. Thus, the complexity of aircraft maintenance basically lies in the individuality of stakeholders, strict maintenance regulations, interactions between stakeholders and the complexity of aircraft structures, etc. The system on the operational support of aircraft maintenance can be recognised as one kind of complex systems.

Agent-Based Modelling and Simulation (ABMS) is an essentially decentralised, individual-centric approach to modelling complex systems composed of interacting, autonomous agents (Helbing 2012). Agents use simple rules and connections to express their behaviours. The behaviours are not explicitly programmed into the models, but arise from agent interactions (Macal and North 2010). The characteristics, involving uncertainty, system dynamics and realism, can be captured by ABMS. Additionally, ABMS supports the modelling the population of objects. Therefore, ABMS is a good choice for modelling the operation of aircraft maintenance, which helps build the relationship between equipment failures and reliability data. In this paper, the ABSS of aircraft maintenance is adapted from the simulation system proposed in Liu et al.

¹<https://www.faa.gov/> (accessed in August 2019)

²<https://www.caa.co.uk/home/> (accessed in August 2019)

(2019).

Fuzzy-Rough Nearest Neighbour (FRNN) is a promising approach for classifying test objects and predicting their decision values. FRNN is principally based on fuzzy set theory, rough set theory and K Nearest Neighbour (KNN) algorithm. Fuzzy set theory (Zadeh 1965) expressed vague information by utilising a class of objects with a continuum of grades of membership, which was used for linguistic representation of patterns, leading to a fuzzy granulation of the feature space. Rough set theory (Pawlak 1982) provided approximations of concepts for the incomplete information, obtaining dependency rules which model informative regions in the granulated feature space. KNN is a well-known technique for the data classification and is one of the simplest approaches for data analysis (Peterson 2009). In ABSS, reliability and fault location information can be considered as condition and decision attributes respectively. The information represents vagueness and incompleteness due to the random process of generating fault location information. Hence, FRNN is a better fit to predict the relationship between reliability and fault location information.

The simulation model is built of practices, rules and data in the real world, which can bring out the interactions between involved objects. Comprehensive knowledge can be learned from simulation data. The knowledge learned from simulation data is capable of defining rules of fault diagnoses. The rules state the way how we are able to find the probably faulty part. The knowledge means the decision table obtained from simulation data. The table gives the relationship between the reliability-related information and the fault location information. The former is obtained by analysing the simulation data. The latter is gained from the stochastic mechanism provided by simulators. In ABSS, each fault happening on one part is always accompanied by its relevant reliability-related information. Hence, the reliability-related information is a quite useful clue in finding the probably faulty part. The FRNN approach can be dedicated to learning the knowledge with the intention of providing rules of fault diagnoses. Therefore, the knowledge of faults enables us to define the rules.

Therefore, in this paper, the major idea of the approach proposed is firstly illustrated in a framework, aiming at integrating the FRNN approach into ABSS for the FCP. Two models are then proposed from different abstract levels, which lay the foundation for implementing the approach. The next step is to technologically accomplish the integration. Finally, simulation experiments are performed to demonstrate the feasibility of the approach proposed.

To the best of our knowledge, it is the first attempt to combine ABSS with the fuzzy-rough sets theory. The approach proposed provides an alternative to improve the fault diagnosis for the aircraft of new types and to accomplish the automatic self-improvement for aircraft maintenance.

The remainder of this paper is organised as follows. Section 2 provides a comprehensive literature review on the FCP for aircraft maintenance and introduces the concepts of fuzzy-rough theory and agent-based simulation. Section 3 proposes a conceptual model and simulation model for combining these two approaches. In Section 4, an approach is designed to implement the FRNN approach in ABSS. The experiments are conducted to analyse the impacts of different strategies on maintenance costs and service level in Section 5. Finally, Section 6 concludes the paper with future perspectives.

2. Literature Review

2.1. *Fault diagnosis for aircraft maintenance*

This subsection provides a comprehensive literature review on fault diagnoses for aircraft maintenance. The scope of researches is limited to the domain of aircraft maintenance. A synthesis analysis of reviewed papers is summarised in table 1. Several aspects are proposed to review the papers, which include proposes simulation models, provides an automated analysis approach, model purposes, and relevance to our research. Moreover, the relevance is evaluated with different levels: low, medium and high.

According to this table, researches seldom simultaneously consider simulation models and fault diagnosis approaches. Fortunately, papers 1 and 3 (Sobie, Freitas, and Nicolai 2018; Wang et al. 2018) get these two approaches together. However, they do not provide an automated analysis method. On the other hand, papers 4 and 8 (Naderi and Khorasani 2018; Bateman, Noura, and Ouladsine 2011) are involved with simulation models, but the models are employed to evaluate the feasibility of approaches. Obviously, no one provides an automated analysis approach on fault diagnoses for aircraft maintenance. Researchers tend to analyse all kinds of fault reports before performing the analysis of fault diagnoses, which causes their approaches partially automated. This is because the reports of natural languages or information redundancy make it hard to easily extract useful information. A wide variety of approaches have been used for fault diagnoses. This also shows that this area is becoming more and more attracting. As for model purposes, most of researches focus on the fault diagnosis for aircraft engines. This is not only because the engine is the most complex component for aircraft but it generates lots of information. The last aspect is associated with the relevance to our research. Few resources from the literature can be reused as no one combines ABSS with the FRNN approach. This is the reason why no paper is given a relevance of high level. Even so, the relevance analysis is still required because others aspects like scopes, advantages, disadvantages etc., can be analysed.

To summarise, the existing works are prone to separating the fault diagnosis from the operating of aircraft, which can cause the analysis of fault diagnoses with insufficient historic data or the data only valid for specific circumstances delayed or impeded. Another difference is that few researchers are devoted to automating the process of the FCP so that fast responses are possible to be made with the aim of dealing with real-time issues on the operational support of aircraft maintenance. Therefore, we are motivated to build a simulation system with the FRNN approach enabling us to automatically support the fault diagnosis for aircraft maintenance without sufficient historic data.

2.2. *Fuzzy-rough theory*

2.2.1. *Fuzzy set theory*

Fuzzy set theory uses given degrees to define the relationship between objects and their memberships. The definitions of fuzzy sets are manifold. In this paper, we introduce the definition of Richard Jensen's (Jensen and Cornelis 2011), which is shown as follows:

A fuzzy set in X is an $X \rightarrow [0,1]$ mapping, while a fuzzy relation in X is a fuzzy

Table 1. An overview of related papers - A (proposes simulation models) and B (provides an automated analysis approach)

article	A	B	applied methods	model purposes	relevance to our research
1.(Sobie, Freitas, and Nicolai 2018)	yes	no	nearest-neighbour dynamics time wrapping	bearing fault classification	medium, similar ideas, no fault prediction, partial automated analysis processes.
2.(Jia et al. 2018)	no	no	complex empirical mode decomposition and random forest theory	fault diagnosis for large wind turbines	low, similar in scope, accurate fault information extraction.
3.(Wang et al. 2018)	yes	no	particle swarm optimisation and back propagation neural network	fault diagnosis for aero-engine fuel regulators	medium, similar ideas, similar forms of representing fault information (fault characteristic table), partial automated analysis processes.
4.(Naderi and Khorasani 2018)	yes	no	data-driven fault detection, isolation and estimation method using Markov parameters	fault detection, isolation, and estimation for aircraft gas turbine engine actuators and sensors	medium, similar in scope, well-suited for real applications, similar ways of evaluation, less restrictive assumption, different ideas, partial automated analysis processes.
5.(Jiao et al. 2017)	no	no	empirical decomposition and probabilistic neural networks	fault diagnosis for airborne fuel pumps	low, similar topics, fast convergence, high efficiency and high performance and recognition for diagnosing typical faults, different ideas, no automated analysis process.
6.(Korvesis 2017)	no	no	neural networks and a regression-based approach	developing data-driven model on predictive maintenance for aircraft	medium, similar ins scope, focusing on fault prediction and investigating data from the operational support for aircraft maintenance, different ideas, partial automated analysis processes.
7.(Tayarani-Bathaie, Vanini, and Khorasani 2014)	no	no	dynamic neural networks	fault diagnosis for gas turbine engines	low, similar in scope, concrete measurable variables for studying fault diagnosis schema, different ideas.
8.(Bateman, Noura, and Oulad-sine 2011)	yes	no	building Fault Detection Diagnosis (FDD) system based a nonlinear model of the aircraft	fault diagnosis and Fault-Tolerant Control (FTC) for an unmanned aerial vehicle	medium, similar in scope, concrete fault detection diagnosis system, a joint FDD-FTC strategy, different ideas.
9.(Kim, Choi, and Kim 2008)	no	no	fuzzy-tuning interacting multiple model	fault diagnosis for aircraft actuators	low, similar in scope, smoothing the fault detection process, low probability of false fault detection, different ideas.

set in $X \times X$. For all y in X , R -forest of y is fuzzy set Ry defined by

$$Ry(x) = R(x, y) \quad (1)$$

for all x in X . If R is a reflexive and symmetric fuzzy relation, that is,

$$Ry(x, x) = 1 \quad (2)$$

$$R(x, y) = R(y, x) \quad (3)$$

holds for all x and y in X , then R is called a fuzzy tolerance relation.

If X is finite, the cardinality of A is calculated by

$$|A| = \sum_{x \in X} A(x) \quad (4)$$

2.2.2. Rough set theory

Rough set theory (RST) is a powerful tool for the rule induction from the incomplete data sets and for concisely extracting the information from a domain. Rough set-based data analysis approaches have been widely, successfully employed in fields like medicine, web and text mining, signal and image processing, software engineering, etc. (Pawlak, Polkowski, and Skowron 2008). As RST has been extended and generalised in different ways, in this paper, we use the definition of RST from Jensen and Cornelis (2011) as well, which is explained in the following.

Let (X, Ω) be an information system, where X is a non-empty set of finite objects (the universe) and Ω is a non-empty finite set of attributes such that $a: X \rightarrow V_a$ for every $a \in \Omega$. V_a is the set of values that attribute a may take. With any $B \subseteq \Omega$ there is an associated equivalence relation R_B :

$$R_B = \{(x, y) \in X^2 \mid \forall a \in B, a(x) = a(y)\} \quad (5)$$

If $(x, y) \in R_B$, then x and y are indiscernible by attributes from B . The equivalence classes of the B -indiscernibility relation are denoted as $[x]_B$. Let $A \subseteq X$. A can be approximated using the information contained within B by constructing the B -lower and B -upper approximations of A :

$$R_B \downarrow A = \{x \in X \mid [x]_B \subseteq A\} \quad (6)$$

$$R_B \uparrow A = \{x \in X \mid [x]_B \cap A \neq \emptyset\} \quad (7)$$

Tuple $\langle R_B \downarrow A, R_B \uparrow A \rangle$ is called a rough set.

2.2.3. Fuzzy-rough set theory

Fuzzy-rough set theory benefits from the advantages of fuzzy and rough set theories, where both the membership of different degrees, and the lower and upper approximations are considered. A general definition (Radzikowska and Kerre 2002) is illustrated

in the following:

$$(R \downarrow A)(x) = \inf_{y \in X} L(R(x, y), A(y)) \quad (8)$$

$$(R \uparrow A)(x) = \sup_{y \in X} T(R(x, y), A(y)) \quad (9)$$

where, L and T are an implicator and a t-norm respectively. In this paper, T_M is defined by $T_M(x, y) = \min(x, y)$, where $x, y \in [0, 1]$. L_M is defined by $L_M(x, y) = \max(1 - x, y)$, where $x, y \in [0, 1]$.

Given fuzzy tolerance relation R and fuzzy set A , tuple $\langle R \downarrow A(x), R \uparrow A(x) \rangle$ is called a fuzzy-rough set.

2.3. Agent-based modelling and simulation

ABSS model is generalised as follows (Macal and North 2010) :

- a set of agents including attributes and behaviours;
- a set of agent connections and ways of interactions;
- an environment where agents live.

ABSS is generally applied when the complexity of the system being modelled is beyond what static models or other techniques can fully present (Helbing 2012). The complexity in real systems is shown as follows (Kellner, Madachy, and Raffo 1999):

- system uncertainty and stochasticity;
- system dynamics;
- feedback mechanisms.

The complexity of the autonomous system of aircraft maintenance mainly due to system uncertainty and dynamics. Uncertainty is reflected not only by random mechanisms but by unknown behaviours arisen from agents' interactions. For example, the occurrence of faults on components of aircraft is conformed to a certain probability. The bounds on these uncertainty issues and the implications of potential outcomes should be understood and evaluated. ABSS provides a flexible and useful mechanism to capture these uncertainties (Heppenstall et al. 2011; Moyaux, Chaib-Draa, and D'Amours 2006). According to Oliveira, Lima, and Montevechi (2016); Layeb et al. (2018), this approach enables us to model the simulation environment incorporating uncertainties, and the real stochastic nature of the environment.

In terms of system dynamics, maintenance tasks change over time, including the maintenance sequences regarding faulty aircraft and the scheduling of maintenance resources. Different scenarios will lead to different system behaviours. A dynamic model can actually manage these changes, however, analytic techniques such as dynamic programming can be intractable because of the high level of complexity. Moyaux, Chaib-Draa, and D'Amours (2006) explained that the system complexity made it difficult to know every possible interaction in the system, because the system only had partial control and observability over its environment, and thus, this environment was highly unpredictable. The agent-based decentralisation took this into account by letting each agent continuously coordinate its actions with other agents, instead of making this agent apply a behaviour prescribed at the design time.

Besides, ABMS also allows us to model real-world systems of interests in ways that beyond the capabilities of traditional modelling techniques, such as discrete event

system or system dynamics (Siebers et al. 2010; Liu et al. 2018). As a result, ABMS is chosen to design the autonomous system of aircraft maintenance.

2.4. Fault Classification and Prediction Analysis Framework

The framework for analysing the FCP on aircraft maintenance is proposed in this subsection. Fig. 1 provides a global view about how to combine FRNN and the ABSS, in order to improve maintenance efficiency and decrease maintenance costs. It consists of the ability to:

- specify the motivation about combining the FRNN approach and the simulation;
- specify the approach of introducing the approach into ABSS;
- specify FRNN-ABSS model with components and connections.

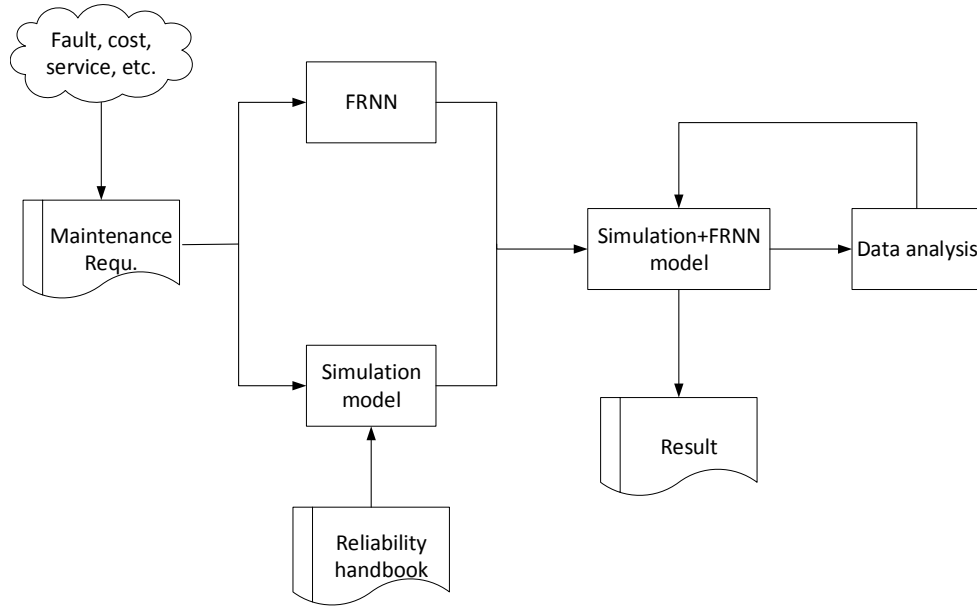


Figure 1. The framework for fault classification and prediction

The FRNN approach is a novel technique to further improve maintenance efficiency from the data point of view, especially when the data shows vagueness and incompleteness. ABSS is able to simulate the whole process for aircraft maintenance, which can produce large amounts of data as well. The data produced by the simulation model can be learned to support decision-makings. Aiming at combing the FRNN approach and ABSS, the reliability handbook (Bombardier. 2004) is used as a basis to propose key indicators. These indicators are a media to connect FRNN and the ABSS model. FRNN-ABSS model is then proposed to realise the improvement of maintenance efficiency from the perspective of data. Since this model is capable of automatically gaining knowledge from the simulation data, the issues of the self-improvement for aircraft maintenance can be performed. Therefore, it is possible for engineers to better define rules of fault diagnoses with the help of the knowledge learned from the simulation data.

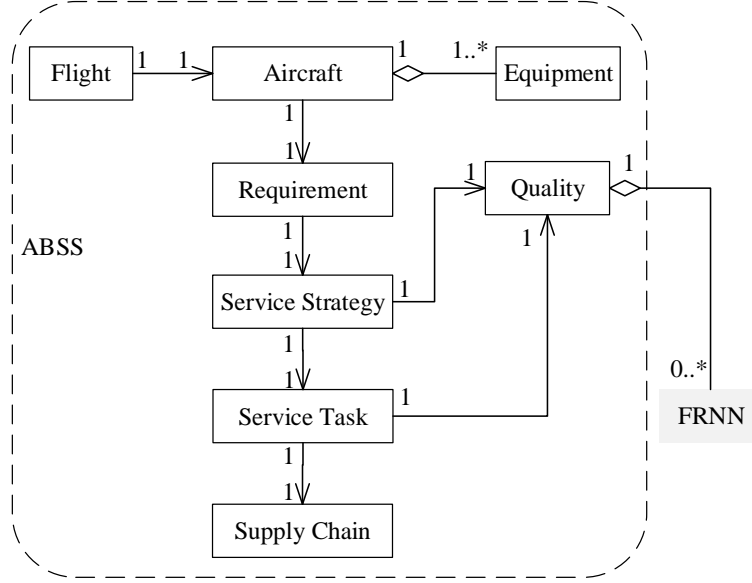


Figure 2. FRNN-ABSS conceptual model

3. Fuzzy-Rough Nearest Neighbour Simulation Model

In order to better describe FRNN-ABSS model, key concepts and their relationships are introduced. Fig. 2 shows FRNN-ABSS conceptual model. Illustrated by this figure, a number of concepts are concerned, including *Flight*, *Aircraft*, *Equipment*, *Requirement*, *Service Strategy*, *Service Task*, *Supply Chain*, *Quality* and *FRNN*. The concepts of *Flight*, *Aircraft* and *Equipment* are the basic elements to describe the maintenance process for aircraft. Concept *Requirement* presents the needs for different maintenance types like scheduled maintenance, unscheduled maintenance, etc., which triggers the maintenance process. The concepts of *Service Strategy*, *Service Task*, *Supply Chain* and *Quality* are associated with maintenance processes. *FRNN* is part of *Quality*. Since the manufacturing of aircraft has been improved significantly, most of the equipment are intelligent enough to alarm themselves. Fault location is much easier than before. However, we are not able to know exactly the functionality of the equipment for the aircraft of new types. We even lack the knowledge on the performance of the combination of equipment. The FRNN approach concentrates on classifying and predicting objects, which enables us to build the relationship between the reliability-related information and equipment failures. Finally, the results from *FRNN* will be synthesised and used as recommendations sent to *Service Strategy* and *Service Task*.

FRNN-ABSS architecture model is shown in Fig. 3. This model is capable of dealing with four kinds of maintenance scenarios: scheduled/unscheduled maintenance and with/without uncertain events. The unscheduled maintenance with uncertain events is determined by fault events and uncertain events. Agents are derived from this model including: Flight, Plane, Equipment Agent, Customer Requirement Agent (CRA), Service Strategy Agent (SSA), Service Task Agent (STA), Supply Chain Agent (SCA), FRNN Agent (FRNNA) and Quality Agent (QA). QA is linked with SSA and STA, in order to collect the reliability-related data. The analysis of FRNNA will aid the decision-making process of the FCP.

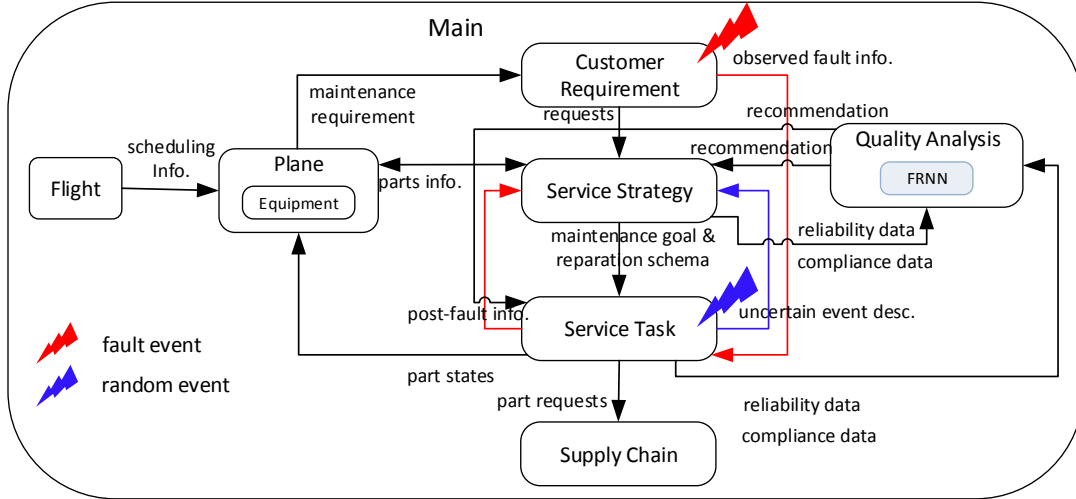


Figure 3. FRNN-ABSS model

4. Fault Classification and Prediction for Aircraft Maintenance

In order to accomplish the FCP of aircraft maintenance, three major issues are considered. The first issue is to understand the relationship between the FRNN approach and ABSS. The FRNN process is the reliability analysis module in ABSS (mentioned in Section 3). ABSS directly made a decision about maintenance issues when the FRNN approach was not considered. Now, ABSS makes a decision after having investigated the prediction results from the FRNN process. Thus, combining the FRNN approach with ABSS should make sure that the latter is able to provide the effective input for the former and the former can deliver the effective output for the latter. The input for the FRNN approach is decision table information and real-time reliability information for parts. The input for ABSS is the specific decision information (which part does not work) regarding parts. The key to combine the FRNN approach with ABSS is the decision table. The basic elements in this table are proposed in Section 4.1. The strategies on the fault analysis of engines are the second issue (Section 4.2). The last one is about how to technologically fulfil the function of the FCP, which is discussed in Section 4.3.

4.1. Reliability-related indicators

The report of reliability on aircraft is a collection of historic information about fault diagnoses, which is employed to analyse faults by associating these experiences with the real-time condition of equipment states. It is issued by aircraft manufacturers and airlines at regular time (Bombardier. 2004), which includes the number of faults on parts, the number of removal parts (scheduled removing and unscheduled removing) on equipment, the flying time and the dispatch times, etc. China Civil Aviation stated that the performance of aircraft can be evaluated by the crew report, the delay of flights, the rate of removing parts, the use time of parts, etc (CCA 2019). Considering the limit of the simulation data collected, several indicators are chosen, including DIR (Departure Interruption Rate), URR (Unscheduled Removal Rate) and MTBUR (Mean Time Between Unscheduled Repair). DIR reflects the delay level of flights. URR and MTBUR are the traditional statistical indicators for reliability analysis.

The reliability office of Shanghai airline provided the definitions of DIR, URR and MTBUR (SA 2019), which are illustrated as follows:

DIR implies the number of delayed flights and cancelled flights (because of technical issues) is divided by the number of all departures during a specific period;

URR indicates the number of unscheduled repairs divided by the flying hours and then multiplied by the number of parts;

MTBUR denotes the amount of use time of parts divided by the number of unscheduled repairs.

Thus, the indicators of DIR, URR and MTBUR will be served as the conditional attributes of the decision table.

4.2. The strategies on the fault analysis for engines

The structure of aircraft is shown in Fig. 4, where the engine component is highlighted. We assume that the warning of the faulty sub-components for the engine can be issued by the controller system of the aircraft. So, the FRNN approach is utilised to predict fault location in sub-components.

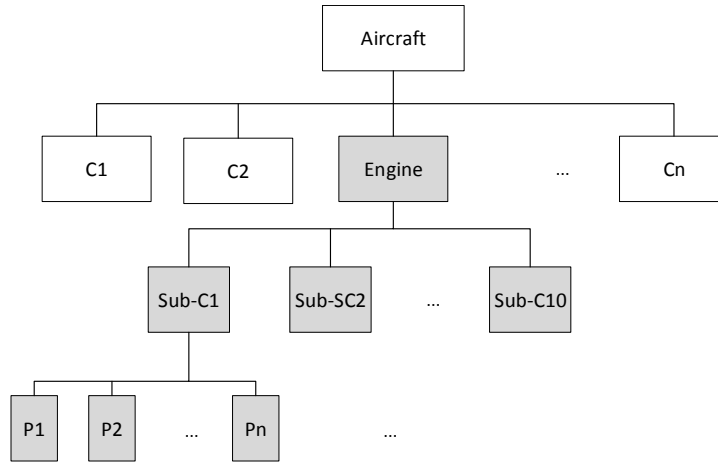


Figure 4. The structure of aircraft

Training sets are the input of the FRNN approach. Two strategies are given for generating training sets, which are shown in Fig. 5. *Random* and *Sequential* strategies are considered, which mean the inspecting process of parts will be executed in a random or a sequential way. The training sets are generated with the help of Algorithm 1.

4.3. Fuzzy-rough nearest neighbour classification and prediction

The objective of the FCP of aircraft maintenance is to find the probably faulty part. Five steps are involved to achieve this aim. First of all, the reliability-related information (DIR, URR, MTBUR) about parts belonging to the faulty component should be collected (Section 4.3.1). The information collected can be treated as the test data. The second step is to calculate the fuzzy tolerance relation between the test data and the training data (knowledge learned from simulation data) with respect to K nearest neighbours (Section 4.3.2). Neighbours are sorted out by calculating the similarities between the instance of the test data and all the instances of the training data. The

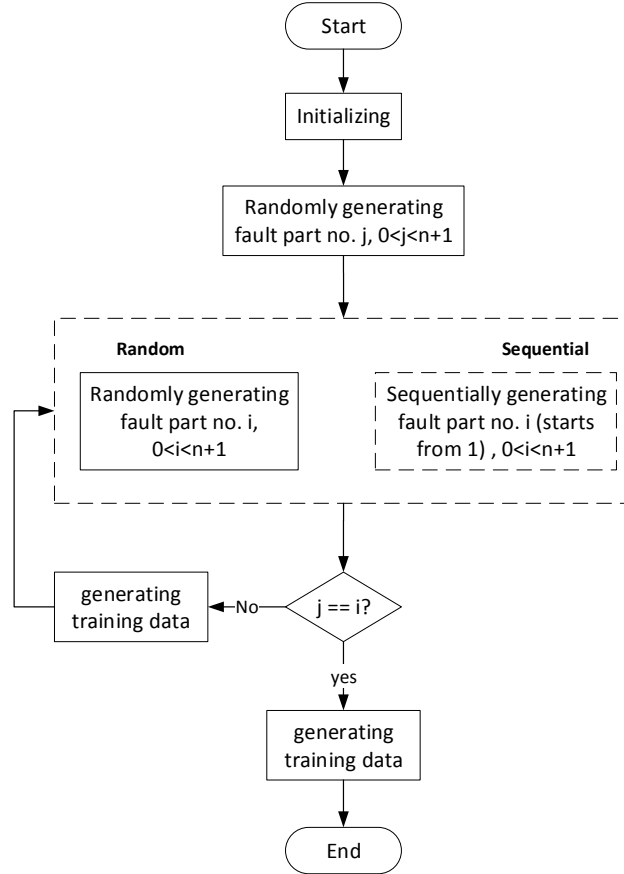


Figure 5. Two strategies for generating training sets

next step is to calculate the lower and upper approximations of one instance of the test data regarding K nearest neighbours (Section 4.3.3). In the following, the averages on the lower and upper approximations of the instance in respect of K nearest neighbours are calculated. The decision attribute corresponding to the biggest average will be assigned to the instance (Section 4.3.4). The last step is to apply the results predicted by the FRNN approach into the process of the FCP for aircraft maintenance (Section 4.3.5). The detailed description of this approach is illustrated in the following.

4.3.1. Decision table generation

The decision table is the input of the FRNN approach, consisting of conditional attributes and decision attributes. It is the media connecting ABSS and the FRNN approach. The conditional attributes were discussed in Section 4.1. The decision attribute is the result of the part inspection, which means it tells whether the inspected part is broken or not.

In order to obtain the decision table, a general process is proposed in Algorithm 1. The major idea is to gather the data of use time, flying hours, departure times, delayed flights and cancelled flights to calculate the values of DIR, URR and MTBUR for relevant parts. This process will also be used to generate test data because it is concerned with these indicators as well.

Algorithm 1: Decision table generation

```
input      : table urr(key, useTime, flyingHours), table dir(key,
              departureTimes)
// The key implies the string of the aircraft type, aircraft no.,
// sub module no. and fault part no.
output    : decision table.

// Collecting raw data for building attributes.
// faultComponent.IsTarget() means the component in which the FRNN
// approach will be used.
1 if faultComponent.IsTarget() then
2   for  $i = 0$  to faultComponent.subComponents.size() do
3     String key = getKey();
4     String useTime = getCurrentUseTime(part(i));
5     String flyingHours = getFlyingHours(part(i));
6     writeToExcel(key, useTime, flyingHours);
7   end
8 end
9 if flight.IsDelayed() || flight.IsCanceled() then
10  if faultComponent.IsTarge() then
11    String key = getKey();
12    String departureTimes = getTotalDepartureTimes(aircraft(i));
13    writeToExcel(key, departureTimes);
14  end
15 end
// Generating the decision table: calculating DIR, URR and MTBUR.
16  $DIR_{part(i)} = count\_tableDir(key) / max\_departureTimes(key)$ ;
// compoentsNumber(key) indicates the number of components
// corresponding to aircraft types
17  $URR_{part(i)} =$ 
    $count\_tableUrr(key) / max\_flyingHours(key) * compoentsNumber(key)$ ;
18  $MTBUR_{part(i)} = max\_useTime(key) / count\_tableUrr(key)$ ;
// DA means the value of decision attributes
19 writeToExcel(dir, urr, mtbur, DA);
```

4.3.2. Similarity calculation

The second step is to calculate the fuzzy tolerance relation R between training data and test data. Jensen and Cornelis (2011) has concluded many options to construct R . Finally, they have chosen the way of calculating R as follows:

$$R(x, y) = \min_{a \in \Omega} R_a(x, y) \quad (10)$$

$$R_a(x, y) = 1 - \left| \frac{a(x) - a(y)}{a_{max} - a_{min}} \right| \quad (11)$$

where, $R_a(x, y)$ is the degree to which objects x and y are similar for attribute a ; a_{max} and a_{min} are the maximal and minimal occurring value of that attribute. Since they have illustrated the feasibility of the FRNN approach for this option, we take this option as well. In our case, Ω includes the conditional attributes of DIR, URR and MTBUR.

4.3.3. Approximation Calculation

The definitions of the lower and upper approximations of a fuzzy set A have been discussed in Section 2.2.3. Here, the definitions of approximations have been adapted from the former ones, where the condition is that the K nearest neighbours have been obtained. The definitions, shown as follows, are the major idea about how to calculate the approximation in this paper.

$$(R \downarrow c)(x) = \inf_{y \in X} \max(1 - R(x, y), A(y)) \quad (12)$$

$$(R \uparrow c)(x) = \sup_{y \in X} \min(R(x, y), A(y)) \quad (13)$$

where c is the decision class; x is the unclassified object (test data); $A(y)$ is to get the value of c corresponding to training data y .

4.3.4. Fuzzy-rough nearest neighbour prediction

The prediction of decision value of test data can be carried out, when the lower and upper approximations have been calculated, The decision class of the maximum mean value of the lower and upper approximations will be chosen as the predictive decision class for test data.

$$c = \max_{c \in C} \frac{(R \downarrow c)(x) + (R \uparrow c)(x)}{2} \quad (14)$$

4.3.5. The fault prediction process for aircraft maintenance

The FRNN prediction process takes the real-time data for parts and generates testing sets, which is illustrated in Fig. 6. Training and testing sets are the inputs of the FRNN approach. The possible set of parts is delivered. The next step is to iterate this set. If the fault part has not been correctly predicted, the rest of the parts will continue to be checked until the right one is chosen.

It should be noted that the above five steps shown in Sections 4.3.1 - 4.3.5 have been implemented as Java code, which allows us to perform the automatic analysis

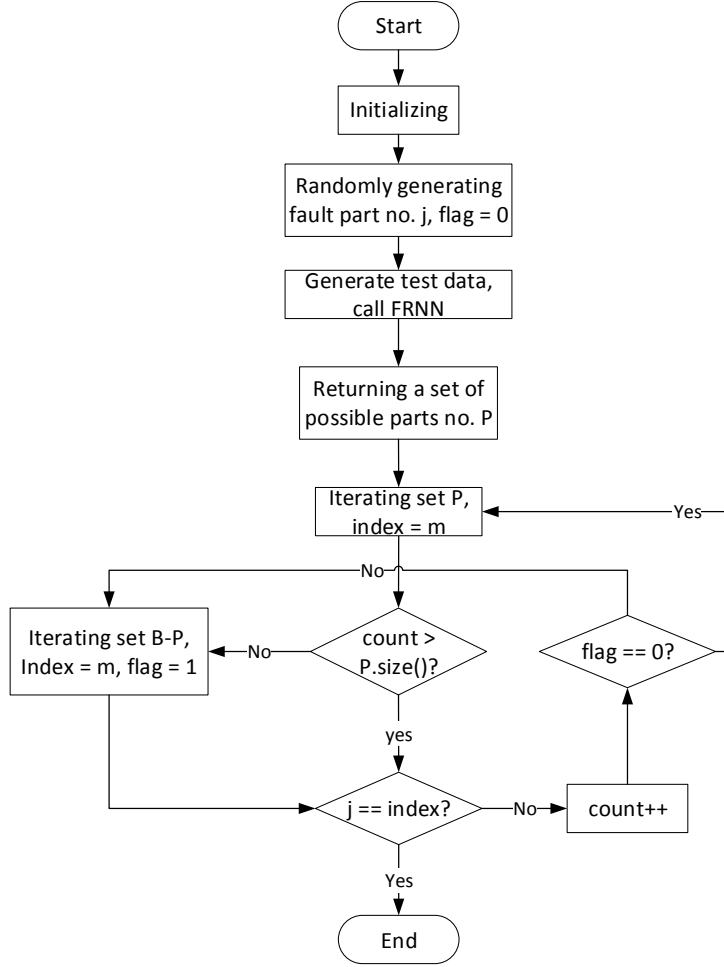


Figure 6. The FRNN prediction process

process on the FCP of aircraft maintenance.

5. Experimentation

5.1. Simulation scenario

The simulation starts from flight schedules. All the aircraft stay at corresponding airports waiting for scheduling signals. The flight legs are collected from the website of *AirFrance*³. Twelve airports are concerned, including *Paris, Lyon, London, Berlin, Madrid, Amsterdam, Boston, Marseille, Barcelona, Prague, Manchester and Munich*. There are 50 aircraft distributed in different airports. Therefore, the aircraft-related stakeholders will be very busy, especially for the maintenance-related. The engine is the critical component in the aircraft, which contains millions of parts. Even though sensors are used to control the major part of the engine, some of components are still not under control. The fault location of engines can waste some of the precious time. Hence, we choose the engine as the target component where the repairing work will

³<https://www.airfrance.fr/> (accessed in November 2018)

involve the FRNN approach, in order to improve the efficiency of the fault location of engines. In this simulation environment, the engine is made of 10 sub-modules containing 73 parts (NASA 2019) (relatively high level of description). Furthermore, this simulation system with the FRNN approach is implemented by Anylogic PLE 8.2.3⁴. The screenshot of this simulation system is shown in Fig. 7. The configurations of the running laptop are 8G memory and 4 processors (i7-6500U CPU).

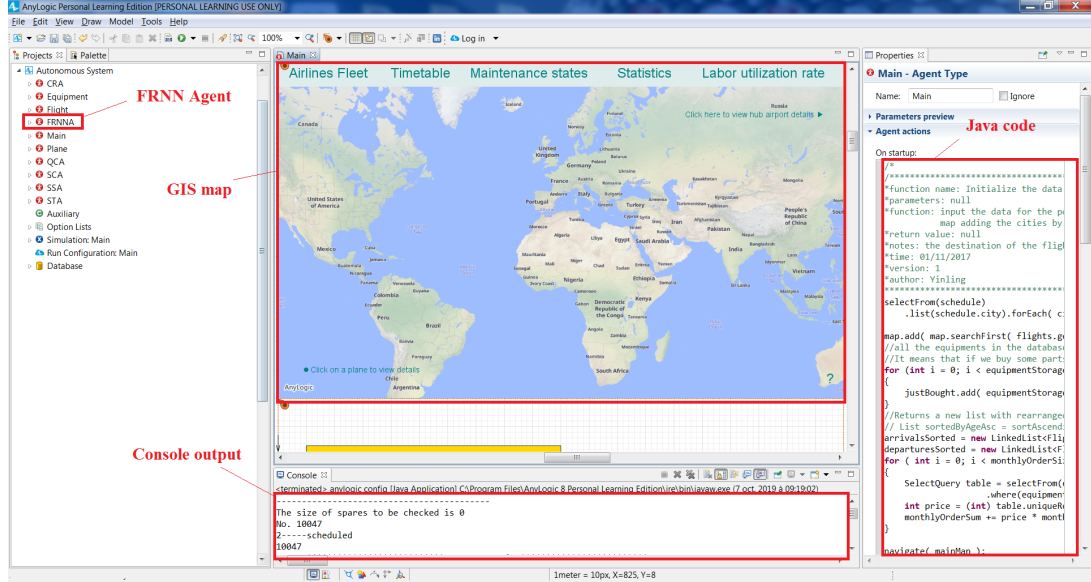


Figure 7. Autonomous system of aircraft maintenance integrating the FRNN approach

5.2. Experiment results

Simulation experiments involve two steps. The first step is to explore the impact of K neighbours on the maintenance cost for different strategies. Optimal K s will then be determined to start the second step which investigates the impact of different strategies on maintenance costs and service level based on the optimal K s.

The impact of the number of neighbours on maintenance efficiency is studied by a set of experiments. The simulation starts from 7:20, 1th January 2017 and stops at 7:20, 1th December 2017, the first 9 months of which are used to generate training sets. The FRNN analysis is performed in the last two months. Here, 15 experiments ($K = 1, 2, \dots, 15$) are conducted for each strategy. The experiment of each K replicates 5 times to avoid chanciness. On top of that the experiments without using the FRNN approach are carried out for comparing the performances of maintenance.

The warm-up period is set to five days to avoid initialisation bias since the simulation system starts with a new environment. The results of the impact of K neighbours on maintenance costs are shown in Fig. 8 and 9 where black broken and red lines mean the experiments are conducted with/without the FRNN approach respectively. The performance of the FRNN approach on *Random* strategy (Fig. 8) is much better than that of *Sequential* strategy (Fig. 9). Fig. 8 implies that the FRNN approach saves the maintenance cost most when the condition is $K = 15$. Fig. 9 shows the choice of K is rather important because more than half of K s lead to lower efficiency of the

⁴<https://www.anylogic.com/> (accessed in November 2018)

FRNN approach. The FRNN approach saves the maintenance cost most when $K = 13$. Therefore, the following analysis on maintenance costs and service level is based on the conditions of $K = 15$ (*Random* strategy) and $K = 13$ (*Sequential* strategy). It should be noted that since our objective is to illustrate the efficiency of this approach and simulation experiments take time (every single simulation experiment takes around 4 hours), we decide to set the maximum of K as 15 instead of continuously increasing the value of K to find the stable performance of the approach.

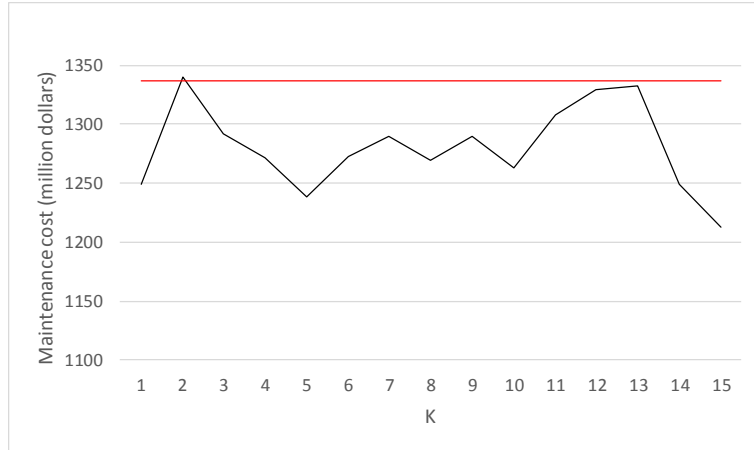


Figure 8. The impact of K nearest neighbours on maintenance costs about *Random* strategy

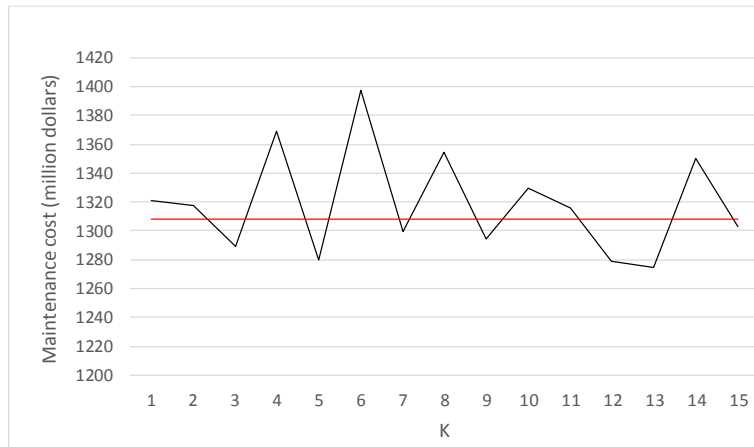


Figure 9. The impact of K nearest neighbours on maintenance costs about *Sequential* strategy

The impact of different strategies (mentioned in Section 4.2) on maintenance costs is illustrated in Fig. 10. The vertical line represents the maximum, average and minimum of 5 replications' data. "+" in the legend means the simulation experiment is executed without using the FRNN approach. This is also applied to Figs. 11 - 13. Fig. 10 shows the averages of *Random* and *Sequential* are lower than those of *Random+* and *Sequential+* respectively. More specifically, 9.3% and 2.5% of the maintenance costs have been saved regarding *Random* and *Sequential* strategies respectively. Thus, the efficiency of the FRNN approach can be observed. The FRNN approach achieves a higher precision with *Random* strategy when comparing with the gaps of differences between averages. The margins of maintenance costs for each strategy suggest that

the experiments of *Random* strategy remain stabler.

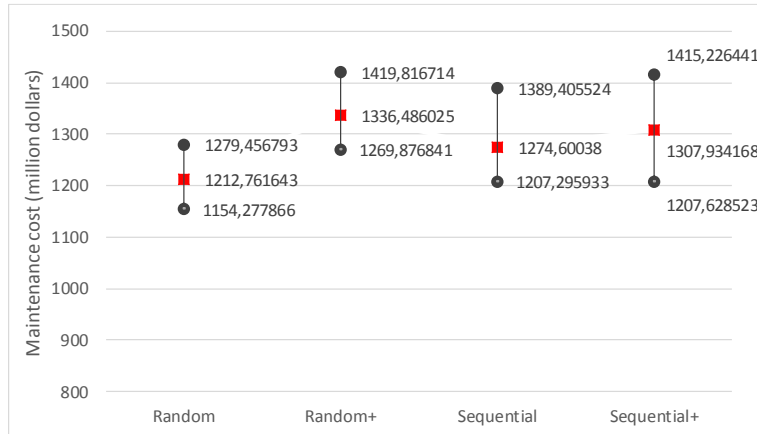


Figure 10. The impact of different strategies on maintenance costs

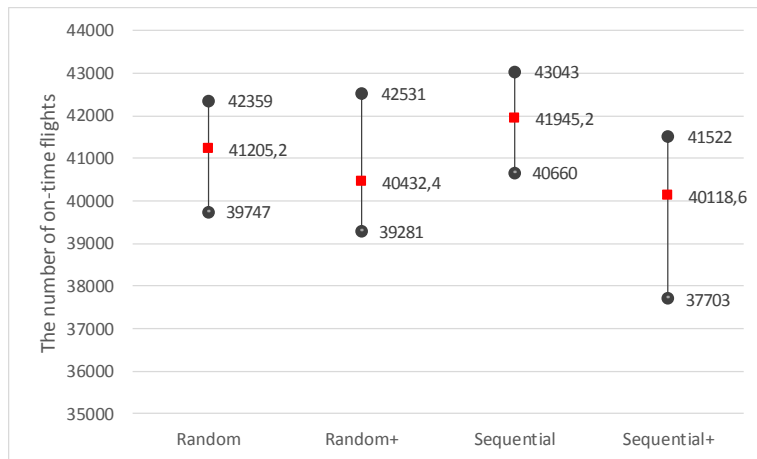


Figure 11. The impact of different strategies on on-time flights

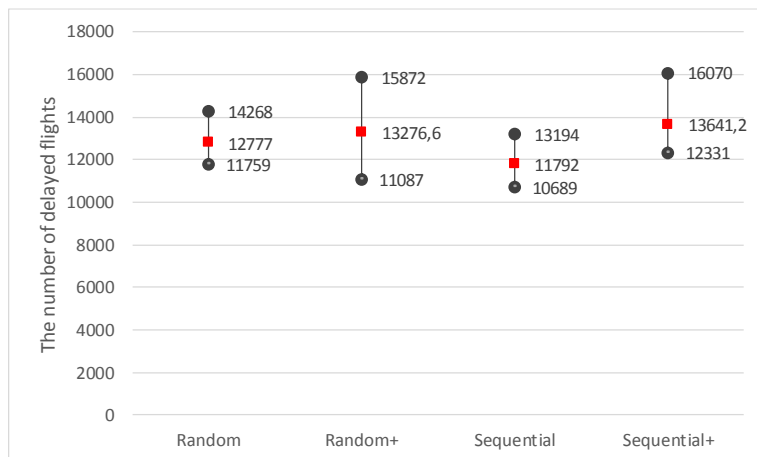


Figure 12. The impact of different strategies on delayed flights

From Figs. 11 - 13, the result of service level is very coherent with that of maintenance costs concerning *Random* strategy. Nevertheless, the result of *Sequential* strategy represents variability. The FRNN approach performs well in terms of on-time flights and delayed flights. As for cancelled flights, the average of *Sequential* is slightly higher than that of *Sequential+*, which depicts the disadvantage of the approach. Even though it is not capable of saving more cancelled flights, its efficiency on *Sequential* strategy still can be observed due to its performance in saving maintenance costs. The downtime costs of maintenance costs imply the difference between total repair times. If the difference of conditions between simulation experiments can be ignored by using averages, the lower maintenance cost means the shorter total repair time. On the other hand, the impact of different strategies on service level in terms of averages is shown in Fig. 14. The result shows that 4.17% and 12.5% of delayed flights have been changed into on-time flights regarding random and sequence strategies respectively. Thus, the FRNN approach also achieves success in predicting faults with respect to *Sequential* strategy.

Service level is another way of examining the efficiency of the FRNN approach. From the above discussion, the analysis of service level helps further investigate the performance of the approach proposed. Therefore, a synthesis of experiment results is necessary to obtain a fair conclusion.

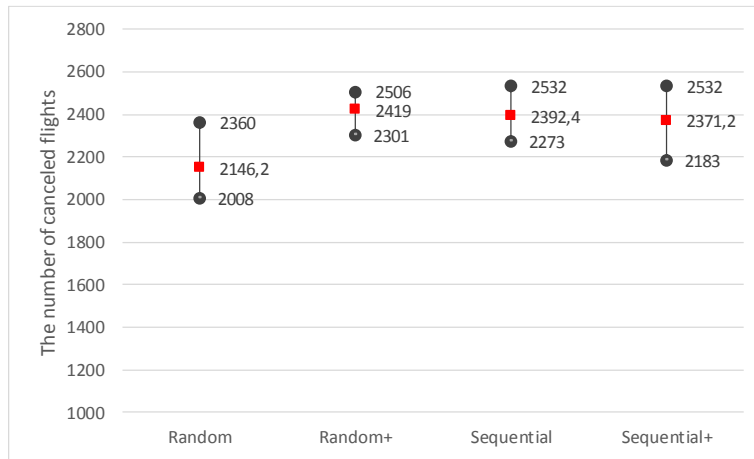


Figure 13. The impact of different strategies on cancelled flights

Above all, we can conclude that the combination of the simulation (ABSS) and the data mining technique (FRNN) can effectively improve maintenance efficiency. The proposed approach makes it possible to investigate the implying fault diagnosis rules for aircraft of new types. To do this, the degrading curves on components of new aircraft are needed to simulate the fault behaviours of components. The simulation system can simulate the life-cycle of aircraft. Thus, a precious data set, including flying hours, inspecting time, repairing time, departure time, delayed flights, cancelled flights, etc., can be collected. In the following, the FRNN approach can be applied to predict the relationship between reliability information and fault location information. Therefore, effective rules of the fault diagnosis can be generated with the virtual experience.

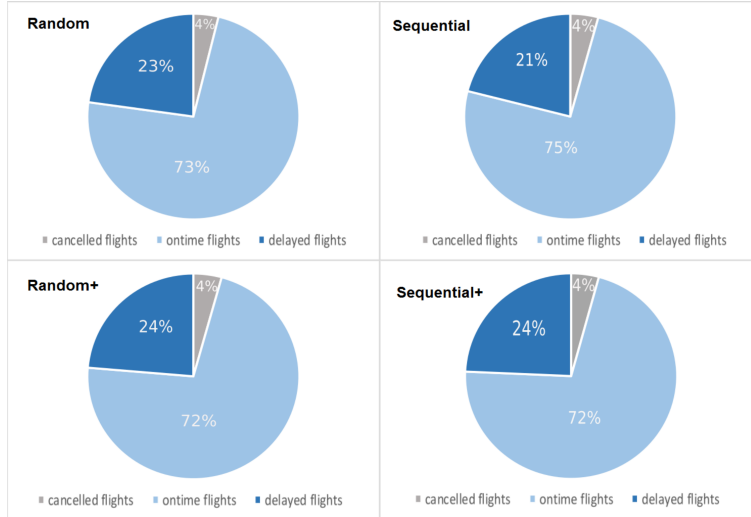


Figure 14. The impact of different strategies on service level in terms of averages

6. Conclusion

In this paper, we discussed the topic about how to accomplish the fault diagnosis for aircraft maintenance without sufficient historic data. ABSS and the FRNN approach were introduced. ABSS simulated the maintenance process for aircraft and produced large amounts of reliability-related data. The FRNN approach performed the FCP based on the data collected. The simulator *Anylogic* was utilised to build the simulation system permitting us to not only simulate the maintenance process but also accomplish the function of the FCP because of the permission of programming in Java into simulation models. That is the reason why we are able to technologically achieve the aim on the automation of the whole maintenance process. The simulation experiments of choosing K nearest neighbours and of analysing random and sequence strategies were conducted to illustrate the feasibility of the approach proposed. Compared with the existing approaches for aircraft maintenance, our approach is the very novel one, which typically holds the following advantages:

- enabling us to perform the data analysis without sufficient real data;
- supporting the automatic self-improvement on the decision-making process for aircraft maintenance;
- a new attempt for analysing the FCP by combing the information system and the data mining.

The feasibility has been demonstrated by experiments, which is able to illustrate the rationality of the maintenance process for aircraft maintenance to some extent. However, the formal verification of maintenance processes is still missing. For example, the verification of issues like Computation Tree Logic (CTL), Linear Temporal Logic (LTL) and deadlock, etc., should be carried out on the model, in order to ensure the consistency and completeness of the model.

NuMSV⁵ is a symbolic model checker originated from the re-engineering, re-implementation and extension of CMU SMV, which allows for the representation of synchronous and asynchronous finite state systems, and for the analysis of specifica-

⁵<http://nusmv.fbk.eu/> (accessed in August 2019)

tions expressed in Computation Tree Logic (CTL) and Linear Temporal Logic (LTL). Therefore, the study of the formal verification of model based on NuSMV will be our future work.

Acknowledgements

This research work was supported by China Scholarship Council.

References

- ASD. 2019. “International specification for developing and continuously improving preventive maintenance.” Accessed 2019-08-10. http://www.s4000p.org/docs/S4000P_Issue_2_0.pdf.
- Bateman, Francois, Hassan Noura, and Mustapha Ouladsine. 2011. “Fault diagnosis and fault-tolerant control strategy for the aerosonde UAV.” *IEEE Transactions on Aerospace and Electronic Systems* 47 (3): 2119–2137.
- Bombardier. 2004. “CSP A2113 CRJ 100 P200.” *Quarterly P Monthly FRACAS report[Z]* 88: 135–136.
- CCA, China Civil Aviation. 2019. “Reliability report for civil aircraft.” <https://wenku.baidu.com/view/3237fcaae43a580216fc700abb68a98271feac07.html>.
- Helbing, Dirk. 2012. “Agent-based modeling.” In *Social self-organization*, 25–70. Springer.
- Heppenstall, Alison J, Andrew T Crooks, Linda M See, and Michael Batty. 2011. *Agent-based models of geographical systems*. Springer Science & Business Media.
- Jensen, Richard, and Chris Cornelis. 2011. “Fuzzy-rough nearest neighbour classification and prediction.” *Theoretical Computer Science* 412 (42): 5871 – 5884.
- Jia, Rong, Fuqi Ma, Jian Dang, Guangyi Liu, and Huizhi Zhang. 2018. “Research on Multidomain Fault Diagnosis of Large Wind Turbines under Complex Environment.” *Complexity* 2018.
- Jiao, Xiaoxuan, Bo Jing, Yifeng Huang, Juan Li, and Guangyue Xu. 2017. “Research on fault diagnosis of airborne fuel pump based on EMD and probabilistic neural networks.” *Microelectronics Reliability* 75: 296–308.
- Kellner, Marc I, Raymond J Madachy, and David M Raffo. 1999. “Software process simulation modeling: why? what? how?” *Journal of Systems and Software* 46 (2-3): 91–105.
- Khoshgoftaar, Taghi M, and Naeem Seliya. 2003. “Fault prediction modeling for software quality estimation: Comparing commonly used techniques.” *Empirical Software Engineering* 8 (3): 255–283.
- Kim, Seungkeun, Jiyoung Choi, and Youdan Kim. 2008. “Fault detection and diagnosis of aircraft actuators using fuzzy-tuning IMM filter.” *IEEE Transactions on Aerospace and Electronic Systems* 44 (3): 940–952.
- Korvesis, Panagiotis. 2017. “Machine Learning for Predictive Maintenance in Aviation.” PhD diss.
- Layeb, Safa Bhar, Amel Jaoua, Asma Jbira, and Yosra Makhoul. 2018. “A Simulation-Optimization Approach for Scheduling in Stochastic Freight Transportation.” *Computers & Industrial Engineering* .
- Liu, Yinling, Tao Wang, Haiqing Zhang, and Vincent Cheutet. 2018. “Information Systems Simulation for Performance Evaluation-Application in Aircraft Maintenance.” In *IFIP International Conference on Product Lifecycle Management*, 789–799. Springer.
- Liu, Yinling, Tao Wang, Haiqing Zhang, Vincent Cheutet, and Guohua Shen. 2019. “The design and simulation of an autonomous system for aircraft maintenance scheduling.” *Computers & Industrial Engineering* 106041.

- Macal, C.M., and M.J. North. 2010. "Tutorial on agent-based modeling and simulation." *Journal of Simulation* 4: 151–162.
- Mechefske, CK, and J Mathew. 1992. "Fault detection and diagnosis in low speed rolling element bearings Part II: The use of nearest neighbour classification." *Mechanical Systems and Signal Processing* 6 (4): 309–316.
- Moyaux, Thierry, Brahim Chaib-Draa, and Sophie D'Amours. 2006. "Supply chain management and multiagent systems: an overview." *Multiagent based supply chain management* 1–27.
- Naderi, Esmacil, and Khashayar Khorasani. 2018. "Data-driven fault detection, isolation and estimation of aircraft gas turbine engine actuator and sensors." *Mechanical Systems and Signal Processing* 100: 415–438.
- NASA, National Aeronautics Space Administration. 2019. "Wright 1903 Engine Parts." <https://www.grc.nasa.gov/www/k-12/airplane/engparts.html>.
- Oliveira, Josenildo Brito, Renato Silva Lima, and José Arnaldo Barra Montevechi. 2016. "Perspectives and relationships in Supply Chain Simulation: A systematic literature review." *Simulation Modelling Practice and Theory* 62: 166–191.
- Pawlak, Zdzisław. 1982. "Rough sets." *International journal of computer & information sciences* 11 (5): 341–356.
- Pawlak, Zdzislaw, Lech Polkowski, and Andrzej Skowron. 2008. "Rough Set Theory." .
- Peterson, Leif E. 2009. "K-nearest neighbor." *Scholarpedia* 4 (2): 1883.
- Radzikowska, Anna Maria, and Etienne E Kerre. 2002. "A comparative study of fuzzy rough sets." *Fuzzy sets and systems* 126 (2): 137–155.
- SA, Shanghai airline. 2019. "Reliability mangement." <https://wenku.baidu.com/view/c78de4e9102de2bd960588bb.html>.
- Sahin, Ferat, M Çetin Yavuz, Ziya Arnavut, and Önder Uluçol. 2007. "Fault diagnosis for airplane engines using Bayesian networks and distributed particle swarm optimization." *Parallel Computing* 33 (2): 124–143.
- Siebers, Peer-Olaf, Charles M Macal, Jeremy Garnett, David Buxton, and Michael Pidd. 2010. "Discrete-event simulation is dead, long live agent-based simulation!" *Journal of Simulation* 4 (3): 204–210.
- Sobie, Cameron, Carina Freitas, and Mike Nicolai. 2018. "Simulation-driven machine learning: Bearing fault classification." *Mechanical Systems and Signal Processing* 99: 403–419.
- Tayarani-Bathaie, S Sina, ZN Sadough Vanini, and Khashayar Khorasani. 2014. "Dynamic neural network-based fault diagnosis of gas turbine engines." *Neurocomputing* 125: 153–165.
- Wang, Ke, Xian Du, Xi-Ming Sun, and Kai Peng. 2018. "Fault Simulation and Diagnosis of the Aero-Engine Fuel Regulator." In *2018 37th Chinese Control Conference (CCC)*, 5783–5789. IEEE.
- Zadeh, L.A. 1965. "Fuzzy sets." *Information and Control* 8: 338–353.