Enhanced Fault Isolation and Part Recommendation for Airplane Health Management with Hybrid Probabilistic Modeling

Partha Adhikari[1], Seema Chopra[2], Darren Macer[3], Dragos Margineantu[4], Surya Pratap Singh Yadav[5] and Thiyagarajan, Sivakumar[6]

[1,2,5,6]*Boeing India Private Limited, Devanahalli, Bengaluru, Karnataka, PIN-562149*
partha.adhikari@boeing.com
seema.chopra@boeing.com
suryapratapsingh.yadav@boeing.com
sivakumar.thiyagarajan@boeing.com

[3,4]*The Boeing Company, 7755 E Marginal Wy S, Seattle, WA 98108, USA*
darren.b.macer@boeing.com
dragos.d.margineantu@boeing.com

ABSTRACT

Aircraft maintenance plays a crucial role in ensuring the safety and reliability of aircraft operations. Effective fault isolation and accurate part recommendation are essential tasks in the maintenance process. The accuracy of existing fault isolation solutions in complex situations (e.g. having multiple fault code scenarios) needs improvement. In this paper, we propose a novel approach of Hybrid Probabilistic Modeling based Fault Isolation Framework combining two solutions. One of the solutions is Pattern Similarity-based Probabilistic Modeling (PSPM) which leverages historical maintenance data to build a probabilistic model that captures patterns of faults and their associated parts replacement. By comparing the current fault symptoms to these patterns, this solution enables more accurate fault isolation and suggests suitable parts for replacement compared to legacy methods. On the other hand, the Physics Informed Probabilistic Modeling (PIPM) employs a Bayesian network to leverage system knowledge in terms of schematics, particularly in scenarios where historical data is sparse or non-existent. Both probabilistic modeling-based solutions complement each other, address gaps, and enhance the efficiency and effectiveness of aircraft fault isolation.

Keywords: Aircraft maintenance, fault isolation, probabilistic modeling, Pattern Similarity-based Probabilistic Modeling, Physics Informed Probabilistic Modeling, Bayesian network.

1. INTRODUCTION

Airplane Health Management (AHM) is a globally accepted strategy to enhance operational availability and reduce maintenance costs across original equipment manufacturers (OEMs), airlines, and operators. Fault diagnostics and prognostics are key steps of AHM. Modern aircraft fault diagnostics techniques have evolved with the installation of more sensors, wider coverage for condition monitoring, and advancement of Built-In Test (BIT). Although the objective for maintainers is to return the aircraft back into service as soon as possible, inadequate means are available for the line maintainer to effectively diagnose from aircraft- or system-level fault to a sub-system component level fault. Preventing effective fault isolation may result in unnecessary part replacements. In scenarios involving co-occurring fault events, existing fault isolation methods often lead to unnecessary part replacements. This results in added costs associated with maintaining a large inventory of spares, multiple attempts to correct an issue, thereby increasing the Aircraft-On-Ground (AOG) time (IATA, 2022).

Several research efforts (Adhikari, 2018) have been made in the field of fault isolation and part recommendation in the context of aircraft maintenance. The following survey of existing work across industries and academia reveals different approaches and techniques used to tackle this problem.

Fault isolation can be broadly categorized into qualitative and quantitative methods. Rule-based, case-based, and graph theory-based are examples of qualitative methods, while Quantitative methods are Model-based (First Principle Physics -based and Data driven model), Machine learning (ML) Classifier-based, Fuzzy Reasoning-, Evidential Reasoning-, and Probabilistic Modeling-based approaches. The approaches for fault-isolation are selected based on the following factors: on-board vs. off-board applications, volume of operational data available, scalability needed, and availability of system safety analysis / design related artifacts.

One prevalent approach is rule-based reasoning (Kramer and Palowitch, 1987 & Ezhilarasu et al., 2019) where a set of rules is established to match observed fault symptoms with known fault patterns. These rules are typically derived from expert knowledge and experience. While rule-based systems can be effective in certain scenarios, they often lack the flexibility to handle complex or novel fault scenarios. Furthermore, developing and maintaining comprehensive rule sets can be labor-intensive, and it requires continuous updates as new fault patterns emerge.

Case-based reasoning (CBR) (Boral et al., 2019 & Deng et al., 2014) solves problems by retrieving similar, previously solved problems and reusing their solutions. The case-based approach requires storing a set of cases where each case holds knowledge about a problem or situation occurring in the past along with the corresponding solution or action. The case-base repository acts as a memory while the act of remembering is achieved using similarity-based retrieval. The retrieved solutions are reused. However, this technique relies only on historical information, is time-consuming and may fail for cases not explicitly represented. The case-based repository grows with each event and can grow very large, making it difficult to manage.

In Graph-based techniques, graphical representation of knowledge (e.g. Bond Graph (Liu and Yu, 2017), Temporal Causal Graph (Mosterman and Biswas, 1997), Timed Failure Propagation Graph: TFPG (Zhang et al., 2020), Multi Signal Dependency Modeling (Li-jia et al., 2018), Minimal Hitting Set (Kleer, 2016 & Pill et al., 2016) and Maximal clique (Bron-Kerbosch-algorithm) are the widely used techniques for domain knowledge representation, but creating the representative model can be time-consuming and it depends on the complexity of the system.

Model-based techniques (Vohnout et al., 2012, Skliros et al., 2021 & Marzat et al., 2013) track nominal system behavior using a first principal physics-based model or a data-driven surrogate model. Then, residuals are generated from the difference between the model's estimated parameters and sensor measurements. Reasoning of residuals and derived features are used for fault isolation.

Physics-based models (Ezhilarasu et al., 2021) require domain expertise, cost, and effort along with having challenges of scalability. Data-driven models (Brown et al., 2007) require a large volume of operational sensor / maintenance data under all possible operational regimes.

Another approach that has gained attention is the use of machine learning-based methods (Ezhilarasu et al., 2021 & Li et al., 2020). These techniques leverage historical maintenance data to train models capable of identifying fault patterns and recommending suitable parts. Supervised learning algorithms such as decision trees, support vector machines, and neural networks have been utilized to classify faults and predict part replacements. Additionally, data mining techniques, including clustering, association rule mining (Yang et al., 2009), and sequential pattern mining (Gao et al., 2013), have been employed to discover hidden patterns and relationships within maintenance data. These approaches aim to identify co-occurrence patterns between faults and parts, uncovering potential associations that may assist in fault isolation and part recommendation. These approaches offer the advantage of automation and data-driven decision-making, but they often require large amounts of labeled data for training, which may be limited in the aircraft maintenance domain.

Probabilistic modeling approaches have also been explored for fault isolation and part recommendation. These methods employ probabilistic graphical models, such as Bayesian networks (Cao et al., 2018, López et al., 2016 & Mengshoel et al., 2013), to represent the dependencies between faults, monitors, and parts. By utilizing historical maintenance data, these models can estimate the likelihood of specific faults and recommend the most probable parts for replacement. Probabilistic models provide a principled framework for reasoning under uncertainty and can handle missing or incomplete data, which is common in aircraft maintenance scenarios.

Despite the advancements in fault isolation and part recommendation techniques, there are still challenges to be addressed. The complexity of aircraft systems, co-occurring fault code situations, variability in maintenance data quality, and the need for efficient decision-making pose ongoing research opportunities. However, the integration of multiple data sources such as sensor data, maintenance logs, and historical fault events can enhance the accuracy and reliability of fault isolation and part recommendation systems.

In this paper, we propose a novel Hybrid Probabilistic modeling-based framework for fault isolation and part recommendation in aircraft maintenance which combines two techniques: Pattern Similarity-based Probabilistic Modeling (PSPM) and Bayesian network-based Physics-Informed Probabilistic Modeling (PIPM). PSPM is applicable when a significant number of similar historical events are available for comparison with the current event, while PIPM plays a role in fault isolation through inferencing a Bayesian network constructed based on the system / sub-system architecture / schematics and logics for fault codes, for cases lacking historical data. This combined framework offers the benefits of both data-driven and model-based solutions.

PSPM looks for similar patterns in historical Maintenance Messages (MMSGs), Flight Deck Effects (FDEs), and associated part replacement data as Final Fix, and it provides enhanced fault isolation and part replacement recommendations. Final fix part is identified when a fault code does not occur over some consecutive flight legs after the fix was done. This paper also presents a fleet level evaluation of a Boeing aircraft model in operation. This paper presents the results of PIPM use in a selection of use cases.

The remainder of this paper is organized as follows: section 2 highlights enhanced fault isolation and Fix Effectiveness overview. Section 3 details the methodology and techniques employed in the PIPM solution with a use case study. Section 4 presents a PSPM overview and fleet level evaluation. Finally, Section 5 concludes the paper and outlines future directions for research in this domain.

2. ENHANCED FAULT ISOLATION AND FIX EFFECTIVENESS OVERVIEW

Our objective is to develop a Probabilistic Modeling-based framework / solution for enhanced fault isolation and part recommendation to perform enhanced fault isolation for co-occurring and single fault code events. Based on the historical maintenance logs, it is observed that unnecessary LRU removal happens during fault isolation, especially in cases of co-occurring (multiple) fault code scenarios.

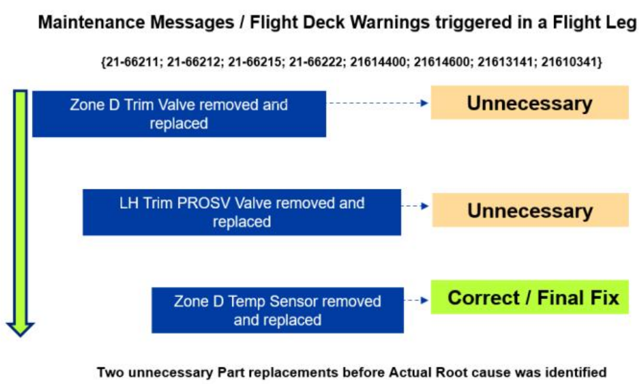

Figure 1: A scenario of occurrence of multiple fault codes and unnecessary part replacement before final fix

Figure 1 contains an example that can explain the existing gap and potential value addition through enhanced fault isolation. Four MMSGs and four FDEs were triggered due to a fault in the Trim Air System of a Boeing aircraft model in a flight leg of a tail of an airline. Two unnecessary replacements happened before the actual root cause or Final Fix was identified. For such multiple fault scenarios, there is a need for improvement in fault isolation due to a higher likelihood of unnecessary maintenance costs and aircraft downtime accrued. Another example can be found in section IV-C (Figure 9). Figure 2 shows the distribution of MMSGs under single and various co-occurring MMSG categories (48.34%) as triggered in historical fleet operation of a Boeing aircraft model. This is the motivation behind this Probabilistic Modeling-based approach, which can improve Fault Isolation / Part Prediction accuracy primarily for the events with co-occurring MMSGs.

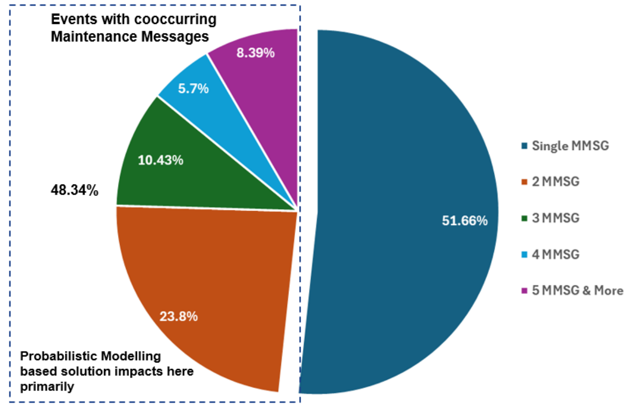

Figure 2: Distribution of triggered MMSGs in historical fleet operation of a Boeing aircraft model.

In the context of AHM, fault isolation is required in two stages (Figure 3): during the early degradation phase, and after the functional failure indicated by triggering of fault codes (e.g., MMSGs and FDEs). In this work, our focus is to isolate the faults after fault codes are triggered (the second phase). However, the same concept can be extended to the first phase: a point prior to the triggering of a fault code such as during precursor anomaly detection (detection of degradation).

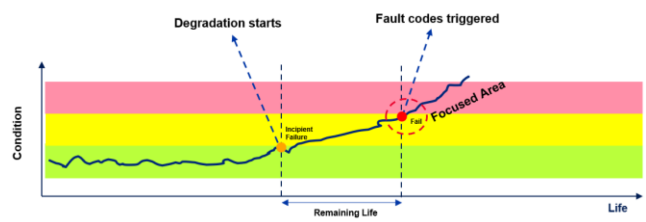

Figure 3: Focus area of fault isolation

The technology vision (Figure 4) for probabilistic modeling-based fault isolation as identified is to combine PSPM and PIPM for enhanced fault isolation

For PSPM, the key enablers are Pattern Mining and Cosine Similarity or Jaccard Similarity. This is a data driven approach applicable when a sufficient number of historical records (at-least 3 to 5) of similar fault code patterns are present with replaced equipment names / numbers for a given pattern. In PIPM, the key enabler is a Bayesian network that embeds schematics of systems, logic for MMSGs, FMEA (Failure Mode and Effect Analysis), etc. within the network model. PIPM is useful when there is a limited number of records of a given pattern of fault codes or when an event that has never occurred in history is observed.

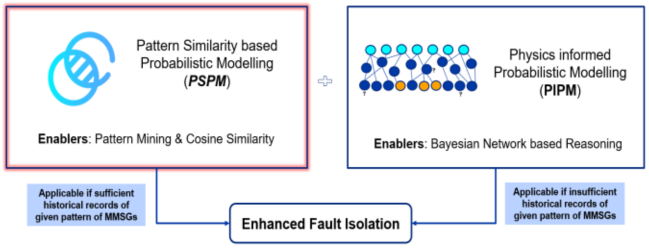

Figure 4: Technology Vision of Probabilistic Modeling based Fault Isolation

Innovation around both methods of Probabilistic Modeling-based fault isolation are carried out. The Bayesian network-based method (PIPM) has been demonstrated with the Trim Air System (Refer to Section III B, Case Study Using PIPM). Further maturity of PIPM and scalability are under study. PSPM is implemented for all Air Transport Association (ATA) chapters of various aircraft models, demonstrating its scalability and key performance indicators of fault isolation. PIPM is described briefly in section III. PSPM is described in more details in section IV.

3. Physics Informed Probabilistic Modeling (PIPM) – Bayesian Network Based

Bayesian networks are at the center of the PIPM based fault isolation solution. This section describes an overview of Bayesian networks (BNs), how to construct scalable BNs from a dependency file (a text file defining logical relationships across nodes of a system along with various properties) defined based on system schematics and other design inputs, as well as reasoning or inferencing of BNs based on given measurements / evidences. A case study of PIPM is presented in this section.

3.1. Overview of PIPM

Bayesian networks or Bayesian belief networks are one of the probabilistic methods used for reasoning under uncertainty. It has been successfully applied in a wide range of domains such as the safety and reliability domains. A Bayesian network (BN) is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG) through a conditional probability table (CPT).

According to conditional independence and the chain rule, BNs represent the joint probability distribution P(X) of variables of any Bayesian network as:

$$P(X) = \prod_{i=1}^{n} p(X_i / parents\ (X_i))$$

BNs can update the prior probability of any event given new information (posterior probability), called evidence M taking advantage of the Bayes theorem:

$$P(X \mid M) = \frac{P(X,M)}{P(M)} = \frac{P(X,M)}{\sum_X P(X,M)}$$

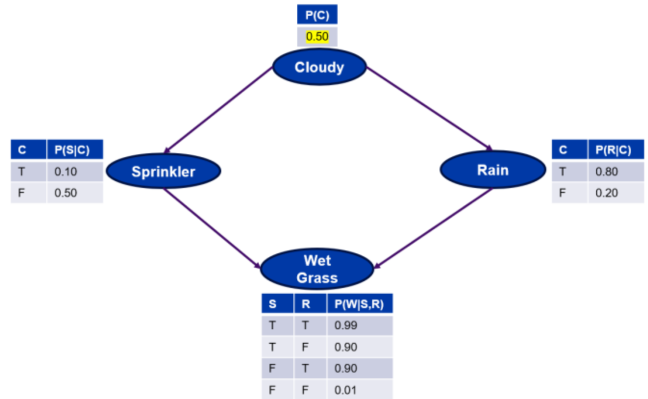

Figure 5: Example a simple Bayesian network (BN)

Let us illustrate a Bayesian network with a simple Grass-Wet-Sprinkler example (Figure 5) where two events can cause the grass to become wet (W): an active sprinkler (S) or rain (R). The presence of clouds has a direct effect on the use of the sprinkler and on the possibility of rain. When the sky is cloudy (C), the sprinkler (S) is usually not active. This situation can be modeled with a Bayesian network. Each variable has two possible values: T (for true) and F (for false). Conditional Probability Tables for the respective nodes are defined based on historical data / experience. Once the Bayesian network is constructed, inferencing / reasoning can be done with known values of various nodes. i.e. evidences. A Bayesian network specifies a joint distribution in a structured form. The full joint distribution: p(X1, X2,....XN) = Π p(Xi | parents(Xi)), which comes from the graph-structured approximation. Accordingly, from the Grass-Wet-Sprinkler BN structure, the joint probability may be expressed as:

P (C, S, R, W) = P(C) P(S|C) P(R|C) P(W|S, R).

The left-hand side is the probability that all of the following events are true: it is cloudy, the sprinkler is on, it is raining, and the grass is wet (P (C, S, R, W)). This gets computed by multiplying the probability that it is cloudy (P(C)) by that of the sprinkler being on (P(S|C)) or that it is raining given that it is cloudy (P(R|C)), as well as by the probability that the grass is wet, given that the sprinkler is on or that it is raining (P(W|S, R)).

The basic task is, given an observation, to infer the probability of an event. For example, it is cloudy: what is the probability that the grass is wet? We want to compute P(W = T|C = T). ($P(W_T|C_T)$ to simplify notation). Re-writing this request in terms of the joint probability:

$$P(W_T|C_T) = \frac{P(W_T, C_T)}{P(C_T)}$$

The denominator is known: 0.5. The numerator may be expressed as a marginal distribution:

$$\begin{aligned} P(W_T, C_T) &= \sum_S \sum_R P(W_T, S, R, C_T) \\ &= \sum_S \sum_R P(W_T|S,R)P(S|C_T)P(R|C_T)P(C_T) \end{aligned}$$

where the summation is over the variable being T, or being F.

From the simple example $P(C_T)$ has simply been cancelled from the numerator and denominator): $P(W_T|C_T) = 0.99 \times 0.1 \times 0.8 + 0.90 \times 0.1 \times 0.2 + 0.90 \times 0.9 \times 0.8 + 0.00 \times 0.9 \times 0.2 = 0.7452$. If the grass is wet and the sprinkler is off, what is the probability of sky being cloudy? If the sprinkler is on what is the probability of the grass being wet? Reasoning can flow in any direction: bottom-up or top-down. Using this reasoning, the root cause behind the grass being wet can be isolated.

Similarly, we can use BNs for aircraft system fault isolation. By constructing the Bayesian network using the sub-system/system schematics and the logic of MMSGs and FDEs, we can compute Conditional Probability Tables which show the cause-effect relationships across nodes (i.e. components) within the network (a directed acyclic graph). To generate evidence for Bayesian network inference, existing sensor measurements related to the sub-system are discretized into health states (e.g. 'Nominal' or 'Degraded') using statistical/machine learning techniques. Constraint programming-based optimization eliminates conflicting nominal evidences which appears due to very low degradation of parameters. Various modules of the framework are shown in Figure 6.

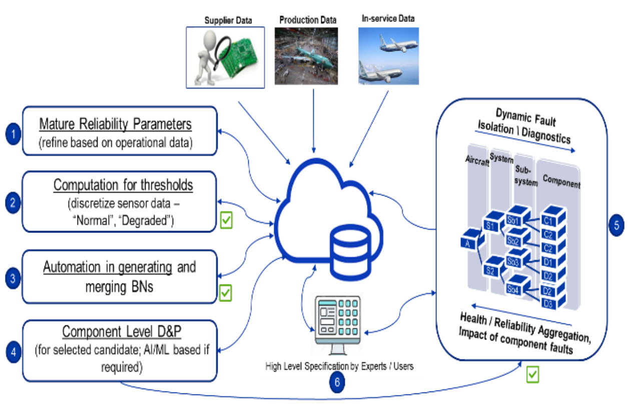

Figure 6: Overview of the Bayesian network-based Fault Isolation Framework

Figure 7 describes the necessary steps to be followed for building the scalable Bayesian network, to inferencing from the BN using discretized sensor measurements as evidence from snapshot data of the Airplane Conditioning and Monitoring System (ACMS) for a given flight. For each subsystem, dependency across various nodes are defined in text form (in Microsoft Excel) from the schematics of sub-systems and logics of MMSGs and FDEs. Then, high-level specifications are merged together to form a system level dependency and a high-level specification file. A Python-based framework has been developed for the automatic generation of the Bayesian network.

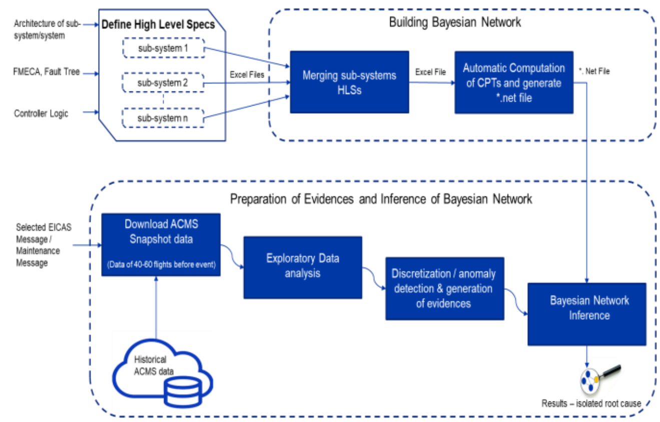

Figure 7: Implementation / Automation of Bayesian network

This module converts high level specifications into a *.net file which defines various node attributes and CPT for each node in a specific format. A CPT is computed from node operation (e.g. 'AND', 'OR', 'NOT', 'EQUAL', etc.) defined in a high-level specification file. As an example (Figure 8), for 'C' node having node operation as 'AND' with parent nodes 'A' and 'B', the computed CPT in specific syntax as per *.net file is defined as: "potential (C | A B) {data = (((1 0)(0 1))((0 1)(0 1)));}". Similarly, for 'E' node having node operation as 'OR' with parent nodes 'C' and 'D', the computed CPT in specific syntax as per *.net file is defined as: "potential (E | C D){ data = (((1 0)(1 0))((1 0)(0 1)));}". Once the Bayesian network is generated, discretized sensor data ("Degraded", "Nominal") and MMSGs and FDEs triggered in the given flight are used for Bayesian network inferencing. Initial validation of inference of constructed Bayesian networks was done using the Java SamIam tool, developed by A. Darwiche's Automated Reasoning Group (UCLA), which is a comprehensive tool for modeling and reasoning with Bayesian networks. A Python based application was developed for the discretization of sensor data, generation of an evidence list and inferencing of BNs for fault isolation using the 'pyAgrum' package.

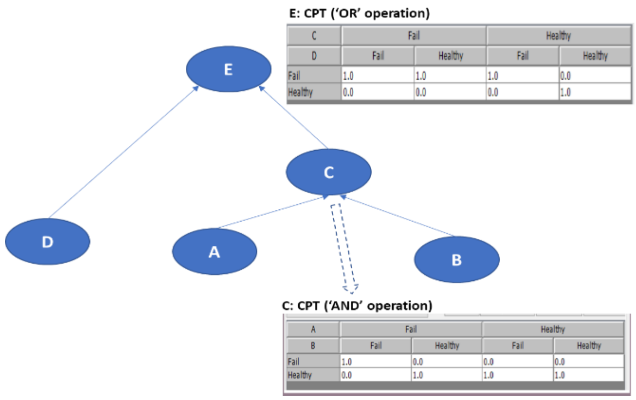

Figure 8: Example of a simple BN with 'OR' node operation (at node C) and 'AND' node operation (at node E)

3.2. Case Study using PIPM

For demonstrating PIPM, the scenario to isolate the valve fault or a sensor fault, as explained in Figure 1 is considered. This is related to the Trim Air System (TAS) of a Cabin Air Compressor Temperature Control System (CACTCS) of a Boeing aircraft model in operation.

The CACTCS consists of Cabin Air Compressors (CAC), Pressurized Air Conditioning Kits (PACK) and Trim Air Systems (TAS). The CAC provides pressurized hot air from which a percentage of air is supplied to the PACK; the remaining percentage gets mixed with cold air from the PACK exit in the TAS before being supplied to cabin zones for air conditioning and cabin pressurization.

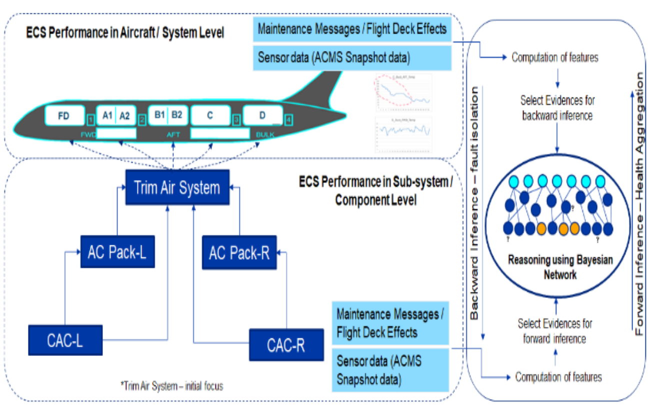

Figure 9: Trim Air system and its interconnection

A Bayesian network was constructed for the Trim Air System. Various steps are shown in Figure 10, where the data in this morphed due to proprietary restrictions. Inferencing of the Bayesian network proved that the zone X AFT duct temperature sensor was degraded for a particular flight. This is matched with the part replacement as mentioned in the maintenance report of the flight leg. Analysis of this scenario using PIPM revealed that this degradation could be detected and isolated almost 22 flight legs in advance.

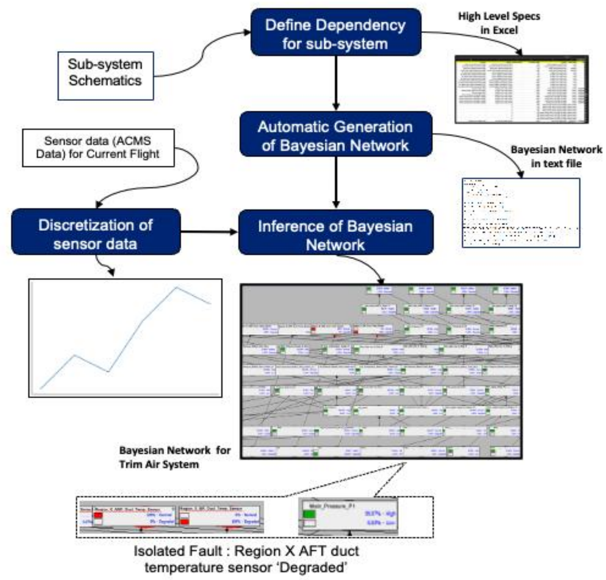

Figure 10: Steps of fault isolation for TAS using PIPM

To construct the Bayesian network, the logic for FDE "TRIM VALVE ZONE X" is used, along with the node dependency based on interconnection of components as shown in Figure 11, where the data in this morphed due to proprietary restrictions.

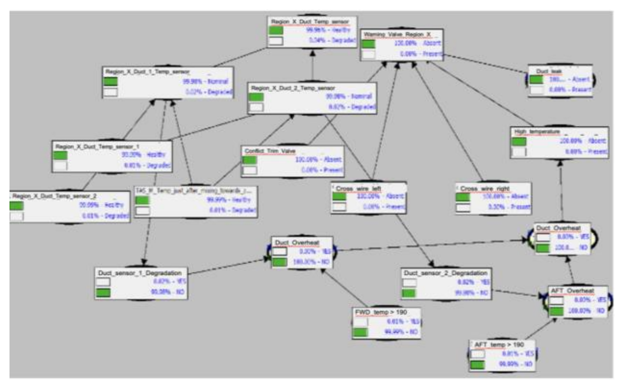

Figure 11: Section of BN with logics of FDEs and dependency based on schematics

Conventionally, Fault Tree or Failure Mode and Effect Analysis (FMEA) are used for various reasoning algorithms including Bayesian networks. In this approach, besides Fault Tree / FMEA, structural dependency across nodes through system schematics and MMSG logics are also used to construct the BN. Though the initial results are promising, there is a need for further research for scalability of BN-based fault isolation solution.

4. PATTERN SIMILARITY BASED PROBABILISTIC MODELING (PSPM)

PSPM underwent various phases of Technology Maturity: concept development, prototype development, implementation, and scalable deployment for various aircraft models and transition to products / services. This section starts with an overview of PSPM followed by fleet level evaluation results.

4.1. Overview of PSPM

PSPM looks for similar patterns in the historical MMSGs and FDEs as well as the associated part replacement data as Final Fix. It then provides enhanced fault isolation and part replacement recommendations. It is to be noted that the PSPM solution is built upon MMSGs, fault codes, and their corresponding maintenance actions. As a result, the solution remains effective regardless of the aircraft's service life. PSPM has two modules: a data pre-processing module (PSPM Module 1), and a part prediction module (PSPM Module 2).

In PSPM Module 1, historical MMSGs and FDEs are downloaded from the Analytics Data Warehouse (A centralized, enterprise-scale repository that consolidates data from many source systems) for the entire fleet of an aircraft model and then unique patterns based on fault codes correlation are identified. Part replacements confirmed as the Final Fix are downloaded from the maintenance logbook. Then it combines these into a single database that maps flight legs, aircraft tails, unique patterns, and associated part replacement data. PSPM Module 1 also has a provision to incorporate sensor data degradation and exceedance patterns across consecutive flight legs.

The PSPM Module 2 uses PSPM DB (computed by PSPM Module 1) and applies a pattern similarity-based approach with adaptive thresholding to recommend the top five parts with their likelihood of failure (which depicts confidence on recommendation) along with maintenance recommendations presenting various historical maintenance actions for similar patterns.

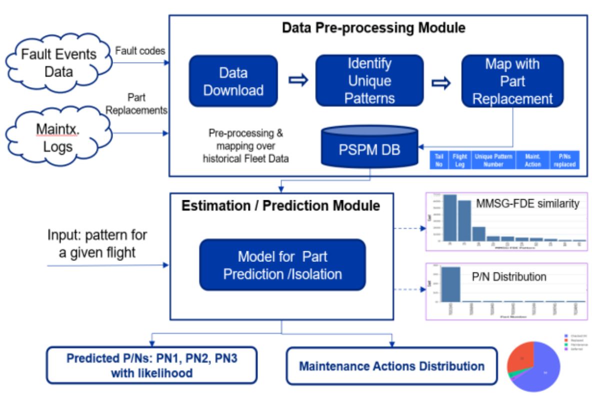

Figure 12: PSPM Functional Block Diagram

Both Cosine (cos(θ)) and Jaccard Similarity (J)- based approaches with adaptive thresholding are implemented. Comparing both approaches it has been observed that the Jaccard Similarity-based approach shows improved performance.

Equations for Cosine and Jaccard Similarity:

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}},$$

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Through the adaptive thresholding logic, PSPM tries to ensure that it provides recommendations to a given pattern with better likelihood. If historical part replacements for similar patterns are found with lesser pattern similarity (< 40%) and part replacement likelihood (< 40%), PSPM does not provide any part replacement recommendation for that pattern. The pattern similarity score & part replacement likelihood are tuned for better results.

The key technical features of PSPM include:

- Identification of unique MMSG and FDE Warning patterns in the entire fleet data for a specific aircraft model and associate them with their maintenance actions along with their part replacement details.
- Utilization of similarity-based algorithms to find the closest match of a given pattern for flight legs exceeding a specified threshold.
- Display of the top five part numbers along with their respective likelihood of failure, the distribution of similar MMSG/FDE patterns, and the historical maintenance actions distribution.
- Adaptive thresholding for automatic selection of the similarity threshold.
- Integration of sensor data for identifying feature exceedance and generating feature trend patterns.
- Implementation of sequential pattern mining (Prefix-projected Sequential Pattern Mining, Peiet et al., 2001) to predict the next probable event in a sequence.
- Inclusion of sensitivity analysis, allowing maintenance personnel to find the distribution of MMSG/FDE patterns for a given part number and display, thereby enhancing decision support for fault isolation and part replacement.

4.2. PSPM Fleet Level Evaluation Study

Fleet level evaluation was done for a single Boeing aircraft model for which the data pool consisted of more than a thousand aircraft tails with flight legs ranging over almost 14 years. All instances of part replacements across every ATA chapter were downloaded for that aircraft model having over 200K fault events. A fleet level evaluation of PSPM solution was performed by taking all the fault events mapped to part replacements. As PSPM is instance-based fault isolation solution, when a fault code pattern triggered in a flight leg (i.e., test event) from the PSPM database is taken for evaluation, all other fault code patterns in the historical fleet represent training set. Similarly, part recommendation was predicted for every fault code pattern in the database, and the following metrics were computed: accuracy, precision, recall and F1 score.

Metrics		Delta for Co-Occurring Fault Events (%)
Top Part Recommendation	Accuracy	5.42
	Precision	4.96
	Recall	5.42
	F1 Score	6.73

Table 1: PSPM Top part Recommendation evaluation results for a Boeing aircraft model at fleet level

Metrics were calculated for the 'Top Part Recommendation' (i.e. the top part recommended matches with the ground truth replacement) and the 'Top 5 Parts Recommendation' (i.e. one of the top five parts recommended matches with the ground truth replacement) for all ATA chapters of a Boeing aircraft model. Table 1 shows the Top part recommendation delta between PSPM's pattern-based approach and the traditional approach of recommending for each individual MMSG. It has been observed that the PSPM demonstrates an improvement of over 5% in part recommendation for pattern-based co-occurring fault events when compared to the traditional method of replacements based on individual fault codes. Table 2 shows the overall PSPM evaluation results (for both cooccurring and single fault events) for Top 5 parts recommendation. With the Top 5 parts recommendation, PSPM achieves prediction accuracy of 96.9%.

Metrics		Scores (%), for both cooccurring & single faults
Top 5 Parts Recommendation	Accuracy	96.9
	Precision	97.09
	Recall	96.9
	F1 Score	96.74

Table 2: PSPM Top 5 parts Recommendation evaluation results of a Boeing aircraft model at Fleet level

Data quality and volume in PSPM pre-processed data as well as emergence of new failure-mode patterns impacts the PSPM part recommendation KPIs. However, if the pre-processed fleet level data is not adequate, PSPM prediction accuracy suffers for some ATA chapters.

5. Conclusion

In this paper two innovative approaches of Probabilistic Modeling based enhanced fault isolation are described with application and results: Pattern Similarity-based Probabilistic Modeling (PSPM) and Physics Informed Probabilistic Modeling (PIPM). The development, implementation, and fleet-level validation of PSPM are discussed in detail, demonstrating its effectiveness in isolating faults in both single and co-occurring fault code scenarios. PSPM has proven to be a robust solution for pattern-based fault diagnosis.

The Bayesian network-based PIPM approach also shows significant potential, particularly in cases involving rare or novel fault events where historical pattern data is limited. However, further research is required to validate the scalability of PIPM, especially in leveraging system schematics and the underlying logic of fault codes and warnings to construct comprehensive probabilistic networks. Additionally, exploring sensor-related pattern analysis and sequential pattern mining will be critical to advancing fault isolation capabilities and improving fix effectiveness. Ultimately, the integration of PIPM with PSPM is expected to yield a more powerful and comprehensive fault isolation framework.

6. Acknowledgements

We acknowledge Eric Nicks, Satish Raghavendran, Brian Pontrello, Mark Mazarek, Franz D Betz, Liessman Sturlaugson, Ameya Kamat, Anusha Pai, Lavanya Konjeti, Varad Mandar Joshi and Pooja Tattur Shivanandappa from Boeing for their unwavering encouragement and invaluable support in various forms.

7. References

Adhikari, P., Rao, H. G., & Buderath, M. (2018, October). *Machine learning based data driven diagnostics & prognostics framework for aircraft predictive maintenance*. In Proceedings of the 10th International Symposium on NDT in Aerospace, Dresden, Germany (pp. 24-26).

Bron-Kerbosch-algorithm, https://en.wikipedia.org/wiki/Bron%E2%80%93Kerbosch_algorithm

C. M. Ezhilarasu, Z. Skaf, and I. K. Jennions (2019). *Progress towards a framework for aerospace vehicle reasoning (FAVER),* in Proc. Annu. Conf. PHM Soc., 2019, vol. 11, no. 1.

C. Skliros, F. Ali, S. King and I. Jennions (Oct. 2021). *Aircraft system-level diagnosis with emphasis on* *maintenance decisions*, volume 236, Issue 6, Journal of Risk and Reliability.

E. R. Brown, N. N. McCollom, E. E. Moore, and A. Hess (January 2007). *Prognostics and Health Management A Data-Driven Approach to Supporting the F-35 Lightning II*, IEEEAC paper #1597, Version 3.

H. Liu and L. Yu (2017). *Analytical Method of Fault Detection and Isolation Based on Bond Graph for Electromechanical Actuator*, IEEE International Conference on Mechatronics and Automation, August.

IATA (2022). *Aircraft Operational Availability*, 2nd Edition

I. Pill, T. Quaritsch, and F. Wotawa (2016). *On the Practical Performance of Minimal Hitting Set Algorithms from a Diagnostic Perspective*, International Journal of Prognostics and Health Management, ISSN2153-2648.

J. Cao, X. Fu, X. Fang, Y. Hu, G. Zhou, and H. Jia (2018) *Bayesian Network based Diagnostics Technique for Civil Aircraft*, IEEE CSAA Guidance, Navigation and Control Conference (GNCC).

J. D. Kleer (2016). *Hitting set algorithms for model-based diagnosis*, 22nd International Workshop on Principles of Diagnostics.

J. Marzat, H Piet-Lahanier, F Damongeot, and E Walter (June 2013). *Control-based fault detection and isolation for autonomous aircraft*, Proc. IMechE vol. 226 Part G: J. Aerospace Engineering.

J. Peiet al., (2001). *Prefix Span: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth*, Intelligent Database Systems Research Lab. School of Computing Science, Simon Fraser University.

J. Zhang, T. Li, C. Chi, and Y. Lv (2020). *A TFPG-based Method of Fault Modeling and Diagnosis for IMA Systems*, Prognostics and Health Management Conference.

L. Li-jia, H. Jian-wang, and S. Hui-xian (2018). *Fault Diagnosis Reasoning Algorithm Based on Multi-signal Model*, MATEC Web of Conferences 173, SMIMA.

M. Li, W. Deng, K. Xiahou, T. Ji, and Q. Wu (2020). *A Data-Driven Method for Fault Detection and Isolation of the Integrated Energy-Based District Heating System*, IEEE Access, volume 8.

M. A. Kramer and B. L. Palowitch (1987). *Rule-Based Approach to Fault Diagnosis Using the Signed Directed Graph*, AIChE Journal 33(7):1067.

O. J. Mengshoel, M. Chaviray, K. Cascioz, S. Poll, A. Darwiche and S. Uckunk (2013). *Efficient Probabilistic Diagnostics for Electrical Power Systems* https://ntrs.nasa.gov/citations/20110012876.

P. J. Mosterman and G. Biswas (1997). *Monitoring, Prediction, and Fault Isolation in Dynamic Physical Systems*, Vanderbilt University, AAAI.

S. Boral, S. K. Chaturvedi, and V.N.A. Naikan (April 2019). *A case-based reasoning system for fault detection and isolation: a case study on complex gearboxes*, Journal of Quality in Maintenance Engineering.

S. Vohnout, B. Kim, N. Kunst, B. Gleeson, R. Wagoner, E. Balaban, and K. Goebel (2012). *A model-based avionic prognostic reasoner (MAPR)*, in Proc. Infotech@Aerospace.

W. Deng, B. Wen, J. Zhou, J. Wang, and Z. Chen (2014). *The Study of Aircraft Fault Diagnosis Method based on the Integration of Case and Rule Reasoning*, IEEE.

Z. Gao, Z. Chen, Y. Feng, and B. Luo (2013). *Mining Sequential Patterns of Predicates for Fault Localization and Understanding*, IEEE 7th International Conference on Software Security and Reliability.

Z. Yang, W. H. Tang, A. Shintemirov, and Q. H. Wu (November 2009). *Association Rule Mining-Based Dissolved Gas Analysis for Fault Diagnosis of Power Transformers*, IEEE Transactions on Systems, Man, and Cybernetics, Part C, volume: 39, Issue: 6.