The Use of Integrated Reasoning with Flight and Historical Maintenance Data to Diagnose Faults and Improve Prognosis

Sylvain Létourneau and Mike Halasz
NRC Institute for Information Technology
1200 Montreal Road
Ottawa, Ontario K1A 0R6
CANADA

sylvain.letourneau@nrc-cnrc.gc.ca / michael.halasz@nrc-cnrc.gc.ca

ABSTRACT

The need for higher aircraft availability and lower maintenance cost is driving the development of Prognostics and Health Management (PHM) technologies. The JSF's Autonomic Logistics (AL) system and the TATEM project are examples of major initiatives that directly rely on PHM for essential enabling technologies. Techniques from artificial intelligence and data mining are expected to provide part of the PHM solution. Research performed at the Institute for Information Technology of the National Research Council of Canada over the last decade demonstrates the usefulness of these techniques for the extraction of knowledge, implementation of PHM reasoning techniques required by decision support tools, and integration of data sources and modelling approaches. However, the process is highly challenging: a significant level of procedural knowledge (or know-how) needs to be developed, the original techniques often need to be extended to achieve an adequate level of accuracy, and the selection of the software development approach appears decisive.

Focusing on the use of existing data from a fleet of commercial aircraft and two PHM applications, this paper illustrates the above difficulties in a very practical manner and introduces corresponding solutions. The first application shows innovative use of artificial intelligence techniques to enhance diagnostics and improve maintenance efficiency at the 1st line. The second application introduces a data mining methodology to build prognostic models from readily available data. For both applications, the usefulness of the proposed solutions in terms of increased availability is discussed. The paper also provides an overview of a generic and open PHM software infrastructure developed to support this research. This software facilitates gradual extensions and integration of PHM techniques.

1.0 INTRODUCTION

Most aircraft operators collect huge amounts of information in central databases. Logistics data, flight data, reliability data, electronic manuals, and aircraft mission data are only a few examples of the information accumulated. Although this data can be effectively used in an isolated manner by staff from various departments, there are a clear lack of solutions that integrate and transform it into actionable knowledge to help organizations achieve their broader objectives. Major initiatives such as the JSF's Autonomic Logistics system [1] and the TATEM project[1] provide leadership in promoting the need for such an integrated solution. In all cases, the anticipated solution requires the development and integration of enabling Prognostics and Health Management (PHM) technologies. Established maintenance methodologies like Reliability Centred

[1] http://www.tatemproject.com/

Maintenance (RCM) also call for continuous evaluation and integration of emerging PHM technologies to help improve diagnostics and prognostics.

Definitions of PHM as found in NAVAIR 00-25-403 and in Hess and al. [1] encompass all devices, analytical methods, and software that can be used for diagnostics, prognostics, health assessment, or health management. For practical reasons, our research follows a progressive integration of theses technologies. Guided by real-world PHM applications, we gradually integrate the techniques that appear most promising. As discussed in Section 5, we have developed a generic open software infrastructure to support this incremental approach. This paper introduces two PHM applications along with the technologies employed. The first application shows innovative use of artificial intelligence techniques to enhance diagnostic and improve maintenance efficiency on the 1st line. The second application focuses on the use of data mining to build prognostic models. For brevity, the paper will only consider applications on commercial aircraft data. Wylie & al. [2] and Yang & Létourneau [3] generalize the concepts presented here and illustrate applications to other kinds of equipment such as mining and railway equipment. The paper is structured as follows. Section 2 describes the context of this research and the data used. Sections 3 and 4 address the two PHM applications mentioned above. Section 5 overviews the open software infrastructure developed to support this research. Section 6 concludes the paper.

2.0 RESEARCH ENVIRONMENT AND DATA

A key mandate of the National Research Council (NRC) of Canada is to help sustain competitiveness of Canadian companies through the development and integration of technology. In 1990, various industries were considered for the application of artificial intelligence (AI) techniques. The maintenance of complex equipment was selected as the preferred target application domain. The economic importance of the maintenance industry in Canada and the expertise of several team members in mechanical engineering warranted this choice. High level economic information available at that point suggested that for every dollar spent on new machinery, an additional 51 cents was also spent on the maintenance of existing equipment [Statistics Canada, 1990]. For 10 industrial sectors, total repair costs exceeded $15 B/year in Canada. Over the following six years, the proportion of the maintenance cost over the acquisition cost increased by 14%.

Initially focusing on the commercial aviation industry, NRC teamed with Air Canada, and GE to study and develop technologies to optimize the use of the available data. For research purposes, the scope was limited to data from Air Canada’s fleet of 70 Airbus A319 and A320 aircraft. These aircraft have systems on board which transmit data to ground stations via Air Canada’s datalink system. This allows monitoring of aircraft health status in a near real-time manner. The data consist of routine performance snapshots (e.g. altitude, temperature, pressures, engine temperatures, valve positions), pilot messages, aircraft generated fault messages, and special purpose reports generated when prescribed limits are exceeded such as on a hard landing. Maintenance data are also available in other systems. They contain descriptions of symptoms and associated maintenance actions in free form text. Some other sources of potentially relevant information could not be made available. These include: deferred problems, flight schedules (static and dynamic), weather, component reliability, check schedules, and parts location. Similarly, information held at the manufacturer, and by people and information systems in the engineering and maintenance control departments were not accessible. The two applications described in the following sections take advantage of different subsets of data.

3.0 USE OF ARTIFICIAL INTELLIGENCE TO IMPROVE DIAGNOSTICS

This section gives an overview of the Integrated Diagnostic System (IDS). Halasz et al [4] provides a more detailed description and discusses the commercialization of the IDS technology. The IDS system is a proof of concept software produced through an early research project on use of AI to optimize maintenance of complex equipment. The objective of the research project was to develop innovative software to improve the efficiency of 1st line maintenance operations at Air Canada. Key functionalities of the proof of concept software developed include: reducing the ambiguity in fault isolation, providing advice on real-time repair action, providing clues on incipient failures, accessing and displaying relevant maintenance information, and facilitation of communication among maintenance staff.

Line technicians in commercial airlines are responsible for maintenance and certification between flights or between legs of a flight. Figure 1 illustrates the typical working environment of the line technician. Their work can be summarized as follows. As soon as an aircraft arrives at the gate, they typically examine pilot’s input (snag messages) and information from the on-board systems for the presence of minimum equipment list (MEL) conditions that could affect aircraft availability. If such conditions exist, they proceed with prescribed troubleshooting procedures and perform repairs as required. Finally, they certify the aircraft for the following flight. This process can involve interaction with other departments of the airline which are either providing resources (e.g., parts and expertise) or waiting for certification of the aircraft in order to proceed with scheduled operations (e.g., system operations and control). Time constraints are typically very severe.

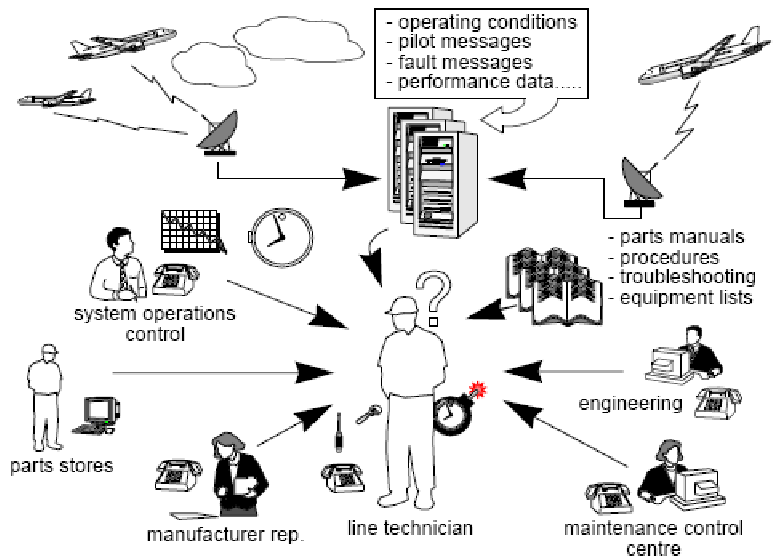

Figure 1: Operational environment around the line technician.

The IDS system supports 1st line activities by integrating and delivering key information in a timely manner to the various staff involved. It also automates some of the steps mentioned above. Specifically, while the aircraft is still in the air, the IDS system: clusters recent warning, failure, and snag (pilot) messages, identifies probable causes by automatically searching the trouble shooting manual (TSM), provides links to the TSM pages as needed, automatically assesses MEL conditions (GO, NOGO, GOIF), and retrieves relevant recent maintenance actions. All the information is displayed in an effective user interface.

The reasoning in IDS relies on two core AI techniques. First, rule-bases are used to encode information contained in the TSM and MEL manuals. The highly structured format of these documents allowed us to develop innovative text processing tools that automatically extract information and generate corresponding rules ready to be implemented in a rule-based system. Rules were also developed to implement heuristics for linking in-flight messages (warning and failure messages) with the TSM, automatically assessing the MEL conditions, and aggregating information with temporal or textual proximity. IDS also exploits case-based reasoning to retrieve relevant historical maintenance information and to suggest potential repairs for the current situation.

To facilitate the evaluation and distribution of the technology developed through this research project, the IDS functionalities were decoupled in a number of well defined modules that can be distributed across an airline's network. Reasoning involved in data fusion, data abstraction, and state assessment is done in server modules. To increase robustness, these server modules can be duplicated as needed. Thin client modules tailored to the needs of the various types of users are deployed on desktops or mobile computers at various sites. Thanks to the reasoning performed by the server modules, only minimal processing is required at the user's device. Figure 2 presents the deployment diagram for the IDS trial. In this case, the server modules were running at the National Research Council in Ottawa while client modules were installed in Montreal and Toronto.

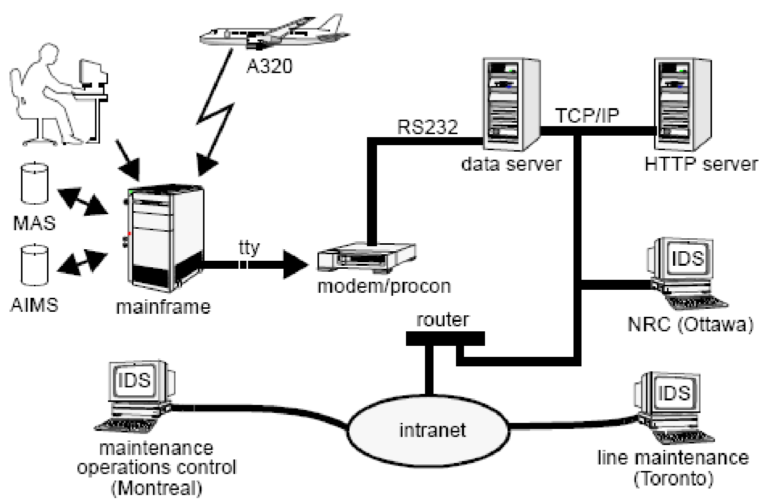

Figure 2: Deployment of the IDS trial system.

Through data fusion and AI-based techniques, the IDS system provided the 1st line technicians with all the information required as early as possible. This information helps technicians maximize their efficiency and it makes it easier to meet the time-constraints imposed for maintenance actions at the gates. This contributes to a reduction in down-time and less disruptions due to unexpected delays. By distributing up to date information on the status of the aircraft to the various staff involved, the IDS system streamlines communications between departments which, in turns, further contributes to increased efficiency. In terms of increased availability, the technology only provides a marginal benefit since it focuses on optimization of operations that happen during relatively short time-windows (turn around times). However, we note that in such time-critical operations even smallest increases in availability can lead to economic benefits.

4.0 USE OF DATA MINING TO IMPROVE PROGNOSTICS

Line technicians are only concerned by imminent maintenance actions. On the other hand, engineering, fleet management, and the part departments would benefit from knowing ahead of time that certain component failures are gradually developing. To extend the IDS concept to these other groups of users, prognostic capabilities need to be integrated. This section briefly discusses a data mining methodology that we are developing to help build the required prognostics models [5].

Figure 3 illustrates the proposed methodology. This methodology builds predictive models for a given component using readily available sensor and maintenance data. There are four steps: data gathering, data transformation, modelling/evaluation, and model fusion. The data gathering step starts by searching the maintenance database to retrieve information on previous replacements of the component of interest. For each case, we record the date of the replacement and the id of the system (e.g., the aircraft tail number or the engine serial number) on which the replacement occurred. In practice, this search may be difficult to automate due to inaccuracies or inconsistencies in the maintenance data. Although, advanced text processing tools could help alleviate these difficulties, manual validation is generally required to ensure that only suitable data is used for modelling. Once the date and the system id for each component replacement have been validated, the data gathering step retrieves sensor data acquired prior to each replacement.

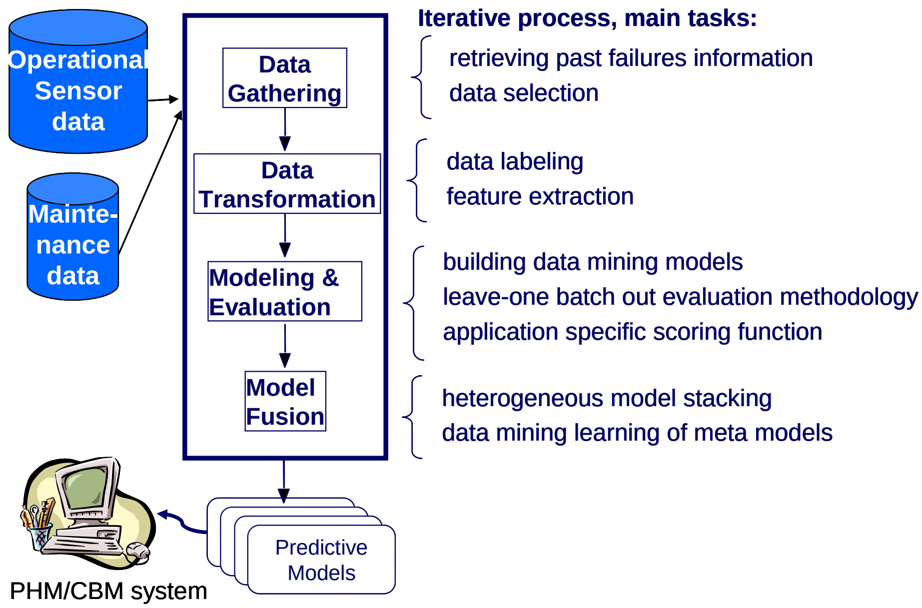

Figure 3: Data mining methodology to build prognostic models.

The data transformation step modifies the obtained sensor dataset in two ways. First, it automatically adds the dependent variable, also named the class or the label. This new variable has two possible values: 1 or 0 for replace component or don't replace component, respectively. The algorithm sets the class to 1 for observations that fall in a pre-determined time interval prior to a replacement (e.g., between 1 to 20 days prior to the replacement) and sets the class to 0 otherwise. Second, several of the initial measurements are normalised to take into consideration the effects of the operating conditions and new time-series variables are included as needed (e.g., FFT, moving average).

In the modelling and evaluation step, data mining and machine learning algorithms are applied to build models that can predict the class as accurately as possible. Several models are created using different algorithms and parameter settings. These models are then evaluated using an evaluation method we have devised to take into account the timeliness of the predictions and the coverage of the models [5].

Finally, if performance improvements are deemed necessary, we complete the process by investigating model fusion approaches [6]. This allows combination of models developed in the previous step instead of selecting the single best one. In some cases, this final step has led to a significant improvement in the quality of the predictions [3].

We applied the proposed data mining methodology in collaborative research projects in aerospace (commercial and military) and to help predict failures with railway equipment. Very promising results have been obtained for building predictive models for components such as starter motors, fuel controllers, and train wheels. On the other hand, research on the application of this methodology to components such as engine's bleed valves did not lead to positive results yet. To help understand the lack of successes with some components, a number of difficulties have been identified including: the lack of relevant measurements, the high level of noise, too many errors in the data, the lack of examples of failures (replacements), inadequate data transformation methods, and the need to integrate more domain information when performing data-driven modelling. Our current research tries to address the last two difficulties through the development of more powerful data transformation methods and the integration of engineering knowledge in the proposed data mining methodology. We are confident that advances in sensor development, signal processing, and an increased awareness of the importance of data quality within the maintenance organization will help resolve the first four difficulties.

Prognostics could help increase aircraft availability by reducing unscheduled down-time, helping to avoid major secondary damage that could ground an aircraft for a long period of time, and minimizing maintenance operations by changing components only when needed instead of at regular intervals. Also, like enhanced diagnostics, reliable prognostics can help reduce disruptions during operation and improve maintenance planning. On the other hand, prognostics usually require sophisticated data acquisition systems which also need maintenance, thus potentially reducing aircraft availability. Accordingly, it is essential to carefully select the candidate components for prognostics and ensure that the maintenance of the data acquisition devices and required software could be done efficiently.

5.0 PHM SOFTWARE ARCHITECTURE

The two applications presented above cover a small portion of PHM technologies. Practical PHM solutions often also integrate technologies from system engineering, materials, and sensors. These technologies are not at the same readiness level and new ones are continuously emerging. An incremental integration approach is therefore mandatory and the software development environment used should facilitate the process. Ideally, the software should also be generic enough to allow the study of PHM applications in various domains (e.g., aerospace, road transportation, railroads). Figure 4 presents a simplified high-level view of a proof of concept software infrastructure we are developing to answer these needs.

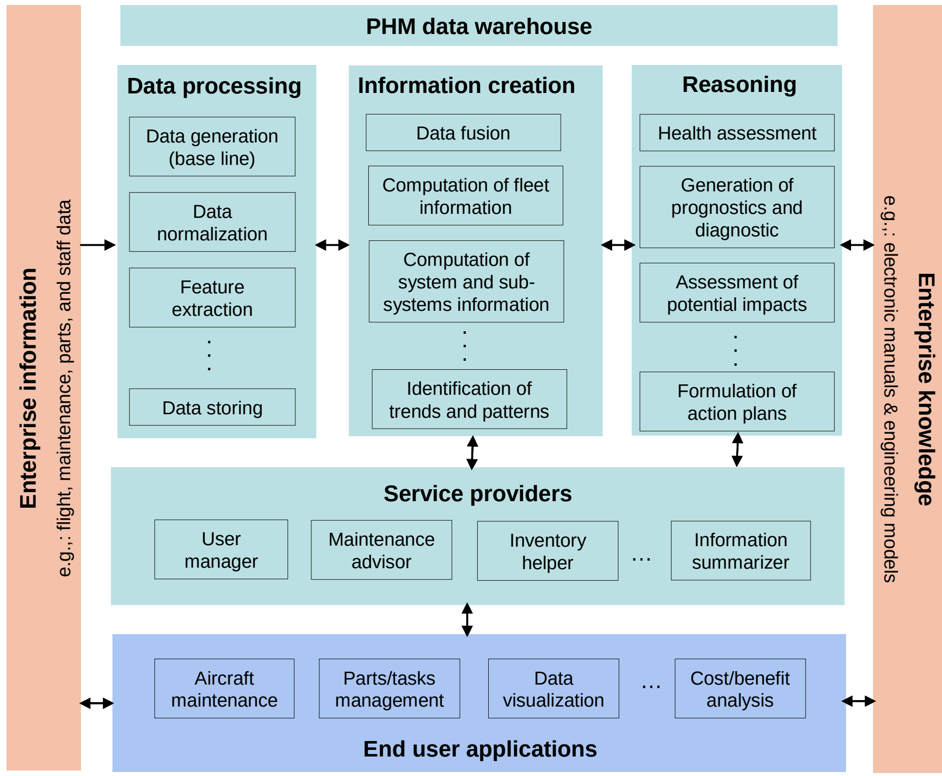

Figure 4: NRC PHM high level infrastructure.

The two modules on the sides represent the data and knowledge available in the organization. Some of the PHM modules can update the organization's information as shown by the bidirectional arrows connected to the side modules. At the top, the `PHM data warehouse` module captures the PHM data. All other modules have access to the information contained in the `PHM data warehouse` but, for simplicity, the diagram does not show the corresponding arrows. The three central modules deliver core PHM functionalities. These are responsible to transform raw information into useful and actionable knowledge. The `service providers` module delivers this knowledge by providing appropriate information to end user applications shown in the module at the bottom. For simplicity, several utility modules have been omitted. These include modules to facilitate communications and data conversion between the various steps.

Each box in Figure 4 may be implemented by one or more algorithms. The system is designed to be robust to addition, modification, and deletion of functionalities. This allows us to gradually integrate new technologies as the R&D progresses. In terms of deployment, the various modules can be distributed over several computers to offer an adequate level of performance or simply installed on a single computer for demonstration purpose.

6.0 CONCLUSIONS

This paper overviewed some of the research performed at the Institute for Information Technology of the National Research Council of Canada to illustrate the potential of artificial intelligence, data mining, and machine learning in the optimization of aircraft maintenance. In particular, we discussed the potential of these technologies to increase aircraft availability. We note that in both real world applications described, the proposed solution integrates various reasoning methods. Moreover, in the case of data mining for prognostics, new methods needed to be developed and integrated in a comprehensive methodology (e.g., automatic labelling and evaluation methods). Since it is generally difficult to select in advanced the techniques to be used and the schema for integration, it is suitable to support the research with generic and open software tools that allow an incremental development of the final solution. The PHM deployment architecture discussed in Section 5 partly answers this need. Current complementary work includes the development of software to support the model development process described in the paper.

7.0 ACKNOWLEDGEMENTS

Many colleagues from the Integrated Reasoning Group[2] of the National Research Council's Institute for Information Technology participated in the research projects described in this paper. Jeff Bird from the National Research Council's Institute for Aerospace Research provided input on the PHM architecture described. Finally, the authors would like to acknowledge the support of Defence Research and Development Canada (DRDC) and Air Canada.

8.0 REFERENCES

[1] A. Hess, G. Calvello, T. Dabney and P. Frith. PHM the Key Enabler for the Joint Strike Fighter (JSF) Autonomic Logistics Support Concept, *Proceedings of the 58th Meeting of the Machinery Failure Prevention Technology (MFPT) Society*. Virginia Beach, Virginia, USA. 2003.

[2] Wylie, R., Mottola, L., Kresta, J., Monk, R. Lessons Learned for the I02 Project, *COM 2002 The Conference of Metalurgists (The International Symposium on Knowledge Management in the Metals Industry)*. Montréal, Québec, Canada. 2002.

[3] Yang, C., Létourneau, S., Learning to Predict Train Wheel Failures, *Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005)*. Chicago, Illinois, USA. 2005.

[4] Halasz, M., Dubé, F., Orchard, R., Ferland, R., The Integrated Diagnostic System (IDS): Remote Monitoring and Decision Support for Commercial Aircraft – Putting Theory into Practice, *Proceedings of AAAI '99 Spring Symposium on AI in Equipment Maintenance*. Palo Alto, California, USA. 1999.

[5] S. Létourneau, C. Yang, C. Drummond, E. Scarlett, J. Valdés, and M. Zaluski, "A Domain Independent Data Mining Methodology for Prognostics", Conference Proceedings: Essential Technologies for Successful Prognostics, 59th Meeting of the Machinery Failure Prevention Technology Society. Virginia Beach, Virginia, USA, April 18-21, 2005.

[6] Lam, L., Classifier Combinations: Implementations and Theoretical Issues, The Proceeding of International Workshop in Multiple Classifier Systems (MCS2000), 77-86

[2] http://iit-iti.nrc-cnrc.gc.ca/about-sujet/ir-ri_e.html