J Intell Inf Syst
DOI 10.1007/s10844-011-0176-1

Aircraft interior failure pattern recognition utilizing text mining and neural networks

Rogério S. Rodrigues · Pedro Paulo Balestrassi · Anderson P. Paiva · Alberto Garcia-Diaz · Fabricio J. Pontes

Received: 11 November 2010 / Revised: 10 August 2011 / Accepted: 11 August 2011

Abstract Being more competitive is routine in the aeronautical sector. Airline competitiveness is affected by such factors as time, price, reliability, availability, safety, technology, quality, and information management. To remain competitive, airlines must promptly identify and correct failures found in their fleet. This study aims at reducing the time spent on identifying and correcting such failures logged. Utilizing Text Mining techniques during the pre-processing phase, our study processes an extensive database of events from commercial regional jets. The result is a unique list of keywords that describes each reported failure. Later, an Artificial Neural Network (ANN) identifies and classifies failure patterns, yielding a respective disposition for

R. S. Rodrigues · P. P. Balestrassi (✉) · A. P. Paiva
Federal University of Itajuba, Itajuba, Brazil
e-mail: pedro@unifei.edu.br

R. S. Rodrigues
e-mail: rogerplisk@gmail.com

A. P. Paiva
e-mail: andersonppaiva@unifei.edu.br

P. P. Balestrassi · A.Garcia-Diaz
University of Tennessee at Knoxville,
401 S Gallaher View Rd #223, Knoxville, TN 37919, USA

A.Garcia-Diaz
e-mail: agarcia2@utk.edu

F. J. Pontes
Universidade Estadual Paulista (UNESP),
Guaratinguetá, SP, Brazil
e-mail: fabricio@embraer.gov.br

Published online: 01 September 2011

a given failure pattern. Approximately five years of historical data was used to build and validate the present model. Results obtained were promising.

Keywords Artificial Neural Network (ANN) · Text mining · Failure pattern · Aircraft log book · Repair

1 Introduction

The aeronautical market faced by airlines is established in a complex environment where people and cargo move in considerable distances around the world. The fierce rivalry ends to settle this market in a scenario where the competitiveness rules. Some reasons can be pointed out as follows:

- Tighten and rigid schedules;
- Pressure from clients that need air freight services;
- Supply time (meals and potable water);
- Maintenance tooling, including technical publications to support maintenance practices;
- Fuel price fluctuation due to speculation and seasonality;
- Crew scheduling (pilots, flight attendants, mechanics, etc.).

In order to analyze this business environment and for better comprehension, the analysis can be divided into different perspectives. Firstly, it can be mentioned the aircraft manufacturers. For each new aircraft project commands considerable amount of technical and economical resources, the minimum details can contribute to a project's success or failure. Such factors are mainly determined during the aircraft conception, but they play an important role for the airlines operation such as:

- Minimum differences during the tender phase for companies that supply aircraft parts can lead in the end to airlines to different maintenance costs (Wang et al. 2008)
- Guaranteed reliability and dispatchability figures offered to the airlines that can later on impact directly on airlines performance (Bineid and Fielding 2003; Kurien et al. 1993);
- Availability of spare parts to support continued airworthiness, low turn-around-time to analyze and repair damaged parts, commitment, and other qualitative or quantitative aspects that sometimes are subjective measured (Farrero et al. 2002).

Also, the aircraft manufacturer has to rely on partners to divide technical and economic responsibilities, just to mention a few. It places the company that plays the partner role in the aeronautical sector in a position which can evidently become even more hostile than of the aircraft manufacturer.

For this, it is not only necessary for partners to offer the most advantageous in order to award a supply contract, but it is also important to offer a good aftermarket support, i.e., to provide a technical support exceeding the customer and aircraft manufacturer's need, with a short timeframe compatible with the sector, thus achieving mutual interests either from airline or aircraft manufacturer perspective.

Knotts (1999) mentions that behind the rigid necessities to comply with the tight itinerary and schedules of the aeronautical sector, an aircraft manufacturer must go beyond the simple task of arranging the competent support. Besides the adversities like heavy air traffic, bad weather conditions, crew and passenger problems, and other unknown circumstances that the airlines have to face, it is also important that the aircraft performance in terms of equipment availability is decisive on determining the purchase of a new aircraft.

In order to have a flight optimized aircraft and airworthy, it is necessary to invest a considerable amount of technical, economical, and logistic resources. These resources shall be allocated by aircraft operators on the best way possible, thus assuring the commitment management (Hansson et al. 2003). Finally, it is possible to keep the aircraft safe and reliable to perform its main function, i.e. transport passengers, crew, and cargo from one point to another, with comfort, safety, and on time.

Now when it is analyzed from the airlines perspective, the practice has shown that the typical airline operational profile seeks a lean operation in terms of costs. To achieve this goal, there is not an absolute truth applicable to all the problems, such the proposed practices found on maintenance manuals, fault and isolation procedures, and so on to have the desired success.

The aircraft operation involves some circumstances such as a considerable number of people and distinct operation conditions. During the aircraft operation, a considerable amount of information is generated. This information can range from different disciplines such as component removal records, log of cabin and system events, and other related flight system information. These data are compiled and then submitted back to the aircraft manufacturer, closing a circle of continuous improvement. Such reports are subject to policies defined by local regulatory authorities of each country, but most of the times the agencies of different countries have a mutual effort to standardize, thus minimizing the necessary time to study the data as well as assuring compatibility and reduced cost. Thus, the aircraft manufacturer or the parts supplier can rely on the field fault finding reported by the airlines, for instance to study and identify trends within the available data.

Although a great effort is done to collect in-service aircraft data, an uncountable number of problems occur. Some problems can be mentioned such as to process and storage of the data (Fernandez et al. 2003), maintenance reports that were incorrectly or improperly filled out due to lack of training of the responsible mechanics, typographical errors, inadequate use of English language, parts erroneously classified, and other problems that can potentially generate a wrong visibility to the manufacturer to solve and attack the problems with less relevance as well as impacting directly on the response time for reported problems. Moreover, this could be also associated to additional hidden costs that are normally unwarranted.

Besides that, the problem is crescent of unstructured manners of storing information among the organizations. Some studies indicate that from 80% to 98% of all electronic data available in the organizations consist of unstructured data, i.e. data that is not available to be easily recovered by means of first hand methods (Wang et al. 2008). In order for the data to be used and manipulated precisely by a computer, they have to be converted into a structured form.

The aircraft systems behavior can be studied in terms of Reliability, Availability, and Maintainability (RAM). Some modeling attempts are mentioned, such as reliability and availability of an aircraft training facility, study of factors direct related to repair such as machinery type, number of machines, age, arrangement of machines, operating conditions, skill level of operating personnel who is in charge of repair the parts, working habits, inter-personnel relationships, absenteeism, environmental conditions, and others. Due to the considerable number of variables, it would be complex to analyze the relationship and interaction of each other in the final impact on RAM (Sekhon et al. 2006).

Sarac (2000) also relates that besides the maintenance operations segment, airlines operations can be based on three other disciplines (Flight Schedule Planning, Fleet Assignment, and Revenue Management). But inside this fertile field, each manufacturer wants to offer to the final customers the best advantage packages as a way to boost aircraft sales, i.e. the airlines will decide to buy aircraft from the manufacturer that will provide the best support ever in the aftermarket. Among the endless advantages that can be offered, one that can be mentioned is that prompt reaction is needed when the failure occurs, to achieve best maintainability practices.

This is closely linked to the business survival. Madu (2000) reinforces the importance of survival of any business enterprise or organization depends on its ability to compete effectively and the necessity of organizations must continuously update themselves in a variety of product and service components that are important to customers and stakeholders.

Finally, this paper presents an analysis of a small portion of this scenario, limiting to the research of failure reports generated during the regular operation of commercial regional aircraft whose parts (from passenger cabin) presented abnormal symptoms and conditions.

2 Background

As per the regulatory authorities, all airlines shall keep records of the events and the related parts for safety reasons. Once an event is reported, an event record must be studied to prevent future failures. These records are returned to the aircraft manufacturer and also to the authority. Each one of them has different responsibilities. The authority may have the concern analyzing the reported events in order to help detect in advance performance and trends. In the United States, the Federal Aviation Administration (FAA) together with other regulatory agencies in the world such as the European Aviation Safety Agency (EASA), Brazilian National Civil Aviation Agency (ANAC), and others instruct that the basic records that should be available are modification status, discrepancies and dispositions, repair listing or identifying the parts used in the repair, maintenance practices used, information about life-limited parts, etc.

The aircraft manufacturer has interest of provide solutions to improve performance, safety, and product robustness. Normally after an event has been detected and the cause (part) has been identified, if possible the operators will try first to make the repair in-house. In the cases that the operators are not able to repair the part in-house, the operators have to identify the vendor where the part came from, either referring to the information available in the maintenance manuals or contacting the aircraft manufacturer who will give the proper guidance on how to proceed with the part subjected to analysis. Sometimes it is under discussion to decide if the operator can claim some compensation due to warranty. Dealing with aircraft parts, they can be either evaluated using the accumulated flight hours or a predetermined timeframe, whichever occurs first. This two-dimensional metric, which was not usual in the past, has been used more often in the recent present, as explained by Chen and Popova (2002).

One of the main consequences that can be enumerated about the reported events in an aircraft is the direct impact on the regular operation, which is measured by the increase of irregularity. Technical delays occur when malfunctioning of equipment and related checking and required corrective action causes the aircraft's departure to be delayed by more than a specified time after the scheduled departure time. Delays are deemed to have occurred if an originating flight departs later than the scheduled departure time, a turn round flight remains on the ground longer than the allowable ground time or if the aircraft is released late from maintenance. Lastly a cancellation occurs if a flight is canceled after being delayed for a longer than expected period of time or due to lack of flight critical and maintainability support.

Although the progress is detected in the sector, there is still space to reduce recurrent costs due to delay. Studies from different authors and agencies show in numbers the impact of delays:

- Knotts (1999) shows that in some cases flight delay costs can reach sums of US$ 775.00 per minute;
- Kumar (1999) discuss the difficulties and losses in planning the resources in airlines.;
- The world's IATA (International Air Transport Association) airlines carried over 800 million passengers in three years (1990–1992), with losses on international scheduled services of US$11.5 billion;
- ATA (2009) estimates that delays in 2007 cost airline customers more than $4 billion in lost productivity and wages. Meanwhile, at a rate of more than $60 per aircraft operating minute, ATA estimates that the 134 million system delay minutes experienced by U.S. airlines in 2007 cost the industry $8.1 billion.

Figures 1 and 2 show respectively the cost of delays per minute in other aircraft types as reference and the main reasons detected that generate such events.

The aforementioned losses presented are also quantified in terms of passenger satisfaction. The local regulation authorities from each country monitors the delay figures closely as a way to measure the airlines performance as well as provide the best services to the population that needs to use the air transportation services. If

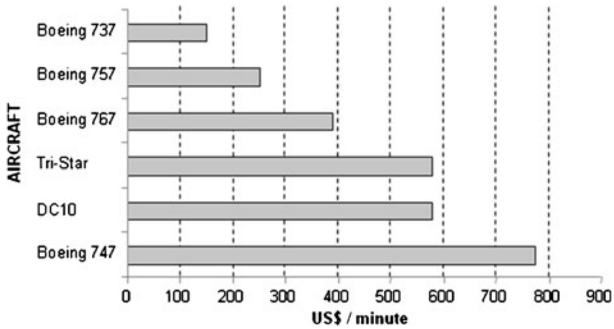


Fig. 1 Delay costs per minute. Reference: Knotts (1999)

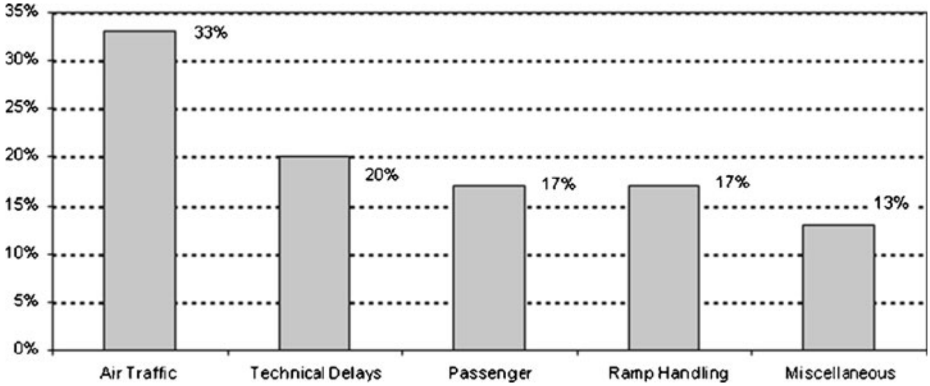


Fig. 2 Delay causes. Reference: Knotts (1999)

the services provided by airlines starts to have a dramatic declination in the figures, consequently it affects the direct satisfaction of the users.

In the United States, where it is based the major market of regional aviation in the world, the BTS (Bureau of Transportation) is responsible to collect and study the numbers of flight departures and flight classified as delayed . It is considered in the U.S. as being the reference for delay situation when a flight departs or arrives more than 15 min behind the scheduled departure/arrival time. They are also quantified by indicators of schedule reliability (SR) and completion rate (CR). These indicators are based on the quantity of delays and interruptions reported by the airline to the aircraft manufacturer. Depending on how it was settled amongst the contracted parties, there are normally penalties if the aircraft performance is not to the agreed values in regards of dispatchability.

The reference is very aggressive when compared to other countries such as Brazil, where it is part of ANAC (National Civil Aviation Agency) to regulate the numbers and parameters for flights that can be classified in delay situations. In Brazil, the

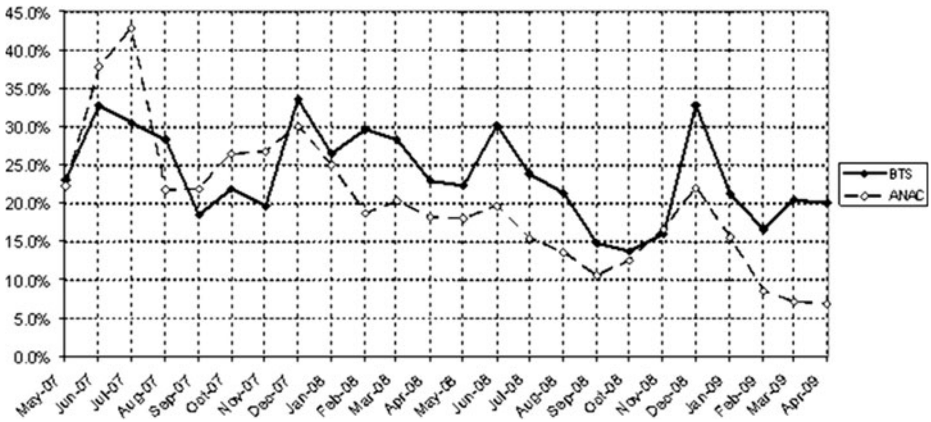


Fig. 3 Delay figures. Reference: BTS (2009) and ANAC (2009)

reference has been changed in the past years as a way to improve the services. Before 2008, it was considered an approximate reference value of 60 min, then after April of 2008 it decreased from 60 to 45 min, and lastly after May of 2008 and onwards the reference is considered as 30 min. To emphasize the importance of spreading the knowledge on solving problems, Knotts (1999) mentions that airlines have a full understanding and great concern of these constraints. When it is not achieved, an undesirable situation of delay is in effect. Figure 3 shows the flight delay percentages for the U.S. and Brazil.

It is interesting to observe that even with different numbers adopted in the countries, the shape of both curves have some similarity. Such similarities can be related to high demand seasons such as holidays and summer, followed by off-peak seasons. To the airlines analysis, the point does not matter as they have always to be vigilant and to study the failure events and the correlation with the degradation of the on time operation.

3 Methodology and literature review

Luxhøj (1999) says that from the perspective of the authorities, aircraft manufacturers, and operators, a good database system has to show—in a convenient time basis—the trends about the failures reported in the field. The roadmap adopted as the methodology to be used in this paper can be summarized by the following steps:

(a) Collection of raw field data;
(b) Populating and pre-processing of raw data utilizing text mining;
(c) Failure pattern recognition utilizing artificial neural networks;
(d) Prediction of future dispositions based on historical data.

The main concepts of text mining and neural networks are described next.

3.1 Text mining

Unstructured text is very common, and in fact may represent the majority of information available to a particular research or data mining project. Wang et al. (2008) discusses the available information in the corporations, where it is normally stored as unstructured text. This unstructured information can be usefully exploited to, for example, identify common clusters of problems and complaints on certain parts, etc.

Prior to the data being introduced into the adaptive system, it was necessary to mine the meaningful data from all the collected reports (Prado and Ferneda 2008). Tan (2008) mentions that with the advent of digital content, databases and archives have received more attention in information retrieval and natural language in the data processing community. Various machine learning processes have been introduced to deal with text classification, such as Centroid Classifier, K-Nearest Neighbor (KNN), Naïve Bayes, Decision Trees, Text classification in Information Retrieval using Winnow, Perceptron, Neural Network. Leitner and Valencia (2008) proposes the combination of human expertise and automatic text-mining systems to gather information from electronically annotated information. There are also techniques to process the natural language, called natural language processing (NLP) discussed by Wang et al. (2008), that have been used successfully to automatically extract useful information from unstructured text using the detailed content analysis on a descriptive text which is what can be used in the present work.

Text mining techniques utilize a combination of one or more state-of-the-art approaches such as mathematics, statistics, natural language processing, and machine learning. For example, Latent Semantic Indexing (LSI) is a mathematic technique that is used in text mining to reduce the dimensionality of text documents (Dumais 2004). Part-of-speech tagging is a natural language processing method that assigns functional labels such as noun, verb, proposition, etc. to words in sentence (Brants 2000). Probability and word frequency are statistical approaches that are commonly used in computing numerical values of terms and in selecting terms. Bayes decision trees (Andrea and Franco 2011) and support vector machines (Song and Chen 2009) are inductive machine learning techniques widely used in text classification. Genetic algorithm (Shamsinejadbabki and Saraee 2011) and unsupervised technique for extracting non-taxonomical relations from domain texts (Punuru and Chen 2011) have also been used.

Text mining has been successfully employed in areas such as Biomedical and Life Sciences, Computer Science, Humanities, Social Sciences and Law, Physics and Astronomy, Business and Economics, etc. There are many interesting fields of research such as detection of similarities between patent documents and scientific publications (Magerman et al. 2010); examining mobile learning trends (Hung and Zhang 2011); discovering a multi-functional metal-binding glycoprotein that exhibits many biological functions of interest to many researchers from the fields of clinical medicine, dentistry, pharmacology, veterinary medicine, nutrition and milk science (Shimazaki and Kushida 2010); identifying fall-related injuries in electronic medical record (Tremblay et al. 2009); mining business policy texts for discovering process models (Li et al. 2010); discovering knowledge by opinion mining from noisy text data (Dey and Haque 2009); tracking what people are saying, finding influencers, and using many social network analytic tools to analyze the underlying social networks embedded within the blogosphere (Macskassy 2011) and (Huang et al. 2011) and with emails via clustering and pattern discovery (Manco et al. 2008); identifying the anomaly cases for knowledge discovery from the warranty and service data in the automotive domain (Rajpathak et al. 2011); discovering frequent musical patterns (motifs) that is a relevant problem in musicology (Jiménez et al. 2011). In Biology, text mining has new challenges as can be seen in Dai et al. (2010); a good example of text mining on language recognition can be seen in Al-Jumaily et al. (2011), where Arabic, the most widely spoken language in the Arab World is identified on the web. Surveys applications of data mining techniques to large text collections, and how those techniques can be used to support the management of science and technology research was importantly explored on Losiewicz et al. (2000).

Related to the present research, Batyrshin and Sheremetov (2008) mention that extracting meaningful data from databases can be useful for management and decision purposes. The development of text mining techniques to extract information from a time series database can contribute on this case. As it is emphasized by the author, the text mining techniques have shown good results to the data owner. The main target is to find among the data the non-suspected relationship in the observed events and then to summarize and categorize effectively and understandably the performed mining. There are several traditional techniques for analysis and data mining, such as segmentation, clustering, classification, indexation, summarization, anomaly detection, notifying discovery, forecasting, and discovery of association rules. Each one has its own particularities and it is suitable towards determining the kind of problem. These tasks can be also mutually related, for instance, the segmentation task can be used for indexing, clustering, summarization, etc. Once the system complexity increases, the sense of making the right and concise decision decreases and up to the moment where the problem is perceptible. The same is applicable to databases where the relevant data are immersed into a memo text field, where it is hard to rely only on the simple human perception to identify possible inherent problems, unless they have reached the necessary significance.

Manning and Schütze (2002) and Bishop (1995) discuss about text mining techniques, where all words found in the input documents and simply count them in order to compute a table of documents and words, i.e., a matrix of frequencies that enumerates the number of times that each word occurs in each document. This basic process is, of course, further refined to exclude certain common words such as "the" and "a" (stop word lists) and to combine different grammatical forms of the same words such as "traveling," "traveled," "travel," etc. (stemming). Similar words can be grouped into clusters, like synonyms. However, once a table of (unique) words (descriptors) by documents has been derived, all standard statistical and data mining techniques can be applied to derive dimensions or clusters of words or documents, or to identify "important" words or terms that best predict another outcome variable of interest.

Once the input documents have been indexed and the initial word frequencies (by document) are computed, a number of additional transformations can be performed to summarize and aggregate the information that was extracted. Many unsupervised feature selection methods have been reported in the literature. The most popular ones are Document Frequency (DF), Log-Frequencies (LF), and Inverse Document Frequency (IDF). Some others are Term Contribution (TC), Term Variance (TV), Information Gain (IG), Mutual Information (MI), $\chi 2$, Relative Document Frequency (RDF), Relative Information Gain (RIG), Relative Mutual Information (RMI) (Prabowo and Thelwall 2006; Yang et al. 2002; Yang and Pedersen 1997) and Neighboring co-occurrence (Huang et al. 2011). All the mentioned methods work in three major steps: (1) Define a formula for measuring the discriminative power of a term. (2) Sort the terms based on the value of defined measurement and (3) Choose a number of the terms from top of the list (Shamsinejadbabki and Saraee 2011).

In this work, after considering many simulations tests, the chosen transformations were:

(a) Log-frequencies:
First, various transformations of the frequency counts can be performed. The raw word or term frequencies generally reflect on how salient or important a word is in each document. Specifically, words that occur with greater frequency in a document are better descriptors of the contents of that document. However, it is not reasonable to assume that the word counts themselves are proportional to their importance as descriptors of the documents. For example, if a word occurs one time in document A, but three times in document B, then it is not necessarily reasonable to conclude that this word is three times as important

a descriptor of document B as compared to document A. Thus, a common transformation of the raw word frequency (*wf*) is to compute:

$$f(wf) = 1 + \log(wf) \text{ for } wf > 0$$

This transformation has a "dampen" effect on the raw frequencies and how they affect the results of subsequent computations. There are other transformations like binary frequencies, where the final matrix has 1 s when the word is found and 0 s to indicate the absence of respective words.

(b) Inverse document frequencies:
Another issue that you may want to consider more carefully and reflect in the indices used in further analyses is the relative document frequencies (*df*) of different words. For example, a term such as "passenger" may occur frequently in all documents, while another term such as "bulkhead" may only occur in a few. The reason is that one might make "passenger" in various contexts, regardless of the specific topic, while "bulkhead" is a more semantically focused term that is only likely to occur in documents that deal with aircraft cabin partitions. A common and very useful transformation that reflects both the specificity of words (document frequencies) as well as the overall frequencies of their occurrences (word frequencies) is the so-called inverse document frequency (for the *i*'th word and *j*'th document):

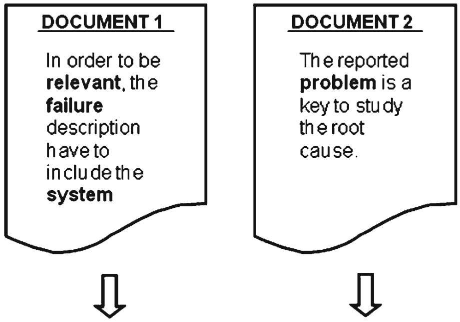
$$idf(i, j) = \begin{cases} 0 & \text{if } wf_{ij} = 0 \\ \left(1 + \log\left(wf_{ij}\right)\right) \log \frac{N}{df_i} & \text{if } wf_{ij} = 1 \end{cases}$$

Since the database is comprised of a set of quantitative and qualitative input and output variables, it is interesting to first extract meaningful data from descriptive fields before applying other tools. Normally a typical database has among the input variables some descriptive fields where important information is stored in an unstructured way. Here the text mining techniques can be useful to reveal hidden pattern points within the descriptive text that could not be possible at a glance for a regular reader.

Tan (2008) shows the concept of vector space model (VSM), where each document is presented as a vector. In this model each document d is considered to be a vector in the term-space. Then, each document D is punctuated similar as cited in Chiang et al. (2008), where the documents D can be converted to a vector representation as a function $d = ((f_1, w_1), (f_2, w_2), \cdots, (f_n, w_n))$, where each f_i represents a document word and w_i represents its frequency (weight). In the context of this work, the document D can be understood by a reported event recorded within a given occurrence.

After the proper transformation is applied to a set of registers, it can be converted to a table where each row represents an event registered and the columns represent each descriptor (keyword) that describes a given event. Figure 4 shows how this transformation is after the text mining applied:

When text mining is implemented, it is possible to convert unstructured (textual) information, extract meaningful numeric indices from the text, and, thus, make the information contained in the text accessible to other learning algorithms such as Artificial Neural Networks (ANN).



		WORDS (DESCRIPTORS)				
DOCUMENT		*relevant*	*failure*	*system*	*problem*	...
D_1		1	1	1	0	...
D_2		0	0	0	1	
...	...	...	...	...	...	...
D_n	...	...	...	...	...	...

Fig. 4 Text mining

3.2 Artificial Neural Networks

ANNs, first used in the fields of cognitive science and engineering, are universal and highly flexible function approximators. As cited by Tsay (2005) apud Balestrassi et al. (2009) ANNs are general and flexible tools for forecasting applications:

> A popular topic in modern data analysis is ANN, which can be classified as a semiparametric method. As opposed to the model-based nonlinear methods, ANNs are data-driven approaches which can capture nonlinear data structures without prior assumption about the underlying relationship in a particular problem.

An ANN, in a simple approximation to human beings, is an attempt of model mathematically the behavior of a biological neural network. Balestrassi et al. (2009) summarizes in a brief manner how a human being brain is composed and how a typical human being neuron is, the response time, and how is the learning process.

An ANN is a computer program that can recognize patterns in a given collection of data and produce a model for that data. It resembles the brain in respect of the knowledge is acquired by the network through a learning process (trial and error) and the interneuron connection strengths known as synaptic weights are used to store the knowledge. The project of an ANN consists of three steps: training, validation, and test.

There are many types of ANN algorithm implemented in many research fields. In this kind of problem, it is common to use the Multilayer Perceptrons (MLP) as a way to have this classification problem modeled. The MLP is the most common form

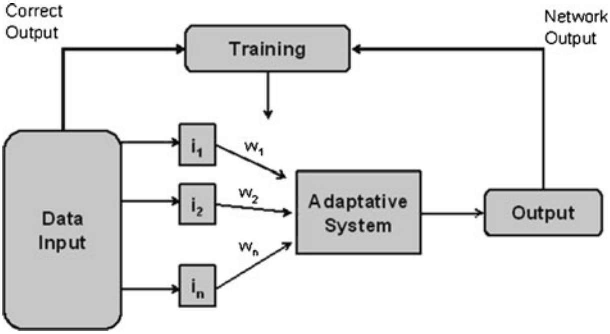


Fig. 5 ANN with backpropagation

of network. Figure 5 illustrates how an ANN using the backpropagation learning algorithm works:

Basically, the training process comprises of presenting data to it. Then, it is computed an output, which is compared to desired output. Sekhon et al. (2006) comments the algorithm provides paired examples of input and output for training the network. Finally, the network weights are modified to reduce error. To use the network, it is necessary to present new data to it, and then the network will compute an output based on its training. As much as cases are available, a better training will be obtained. It is also relevant to resemble that the use of historical data does not mean that all the future cases can be assessed, so that relationships that held in the past may no longer hold. Then, during the validation step, what is tested is the generalization efficiency acquired during the training step. Lastly, the test step is used to do the performance test during the network utilization.

4 The dataset

Basically when an event is observed, it is mandatory that the flight crew registers the event (problem) on the aircraft log book. Then the event is analyzed by the mechanic and an action is carried out, closing the loop. Sometimes the parties do not even know that such data is a valuable collection and an implied feedback from the field, how the product is performing, if the warranty time guaranteed values have been achieved, the lessons learned that would be used for further product development cycles, lastly as varied applications.

It is mandatory that the operators keep the activities tracked, because inspection and diagnostic activities are integral components of an effective maintenance strategy in an attempt to ensure aviation system safety, reliability, and availability.

Although the dataset requirements are clearly stated by the regulatory authorities, the operators may have different ways (different dataset architecture) to store the log books. In Brazil, the National Civil Aviation Agency instructs the airlines as per RBHA (Brazilian Regulations for Aircraft Certification) number 91, subpart E (maintenance, preventive maintenance, modifications, and repairs), section 91.417; RBHA number 121, subpart T (flight operations), section 121.563, subpart V (register and reports) and section 121.705 (summary report for mechanic interruption). In the United States, the analog regulation is the FAR (Federal Aviation Regulation), also known as CFR (Code of Federal Regulation), with the same parameters. These parameters should have at minimum the following information:

1. Aircraft manufacturer, aircraft model, aircraft serial/tail number, engine or propeller;
2. Aircraft registration;
3. Airline name;
4. Date when the failure or defect was identified;
5. What flight phase the failure was identified;
6. What failure or defect identified;
7. Applicable ATA system chapter and subsystem;
8. Total flight hours or flight cycles accumulated by the defective part;
9. Part manufacturer, part number, part description, part serial number;
10. What actions or emergency precautions where performed;
11. Other relevant information that can help to find the failure cause, malfunction or defect, including time since last maintenance action, revision, repair or inspection;

According to ATA (2009), a unified database to track such failures can contribute greatly for reducing administrative time involved in the interchange and processing of provisioning data, procurement transactions, invoices, and consumption data. In addition, the benefits to be gained in utilizing standardization increased and in some cases they have been a prime consideration in the decision process of implementation of ATA iSpec 2000. The idea is to keep the data as simple as possible and at the same time complying with the regulatory requirements. Figure 6 shows how it works:

As the uncountable control carries variables, a repair process can have aside from dates, an Air Transport Association (ATA) reference number (i.e. the ATA

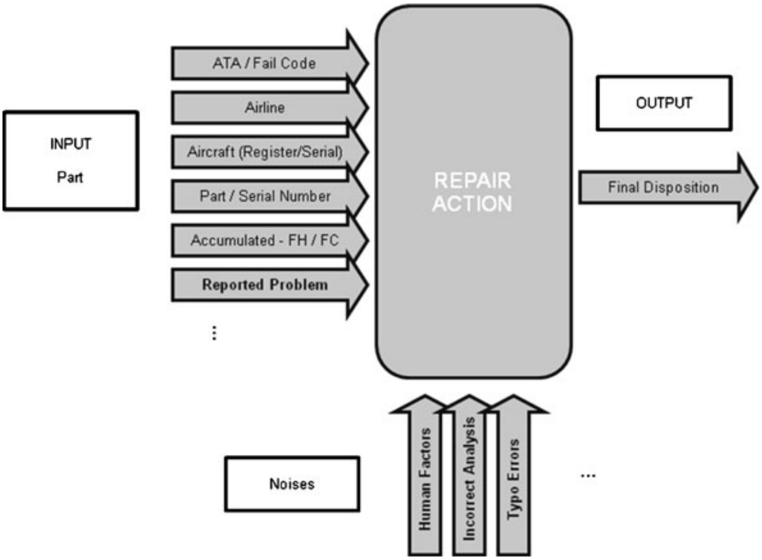


Fig. 6 Reported failures structure

assigns for aircraft parts a number reference, composed of three pair of digits. The pairs stand for the chapter—system, section—subsystem, and subject—unit), name of the operator who sent the part, aircraft data (register number and serial number), part accumulated flight hours or flight cycles (i.e. one cycle corresponds to the action of take off and landing), in-and-out dates, analysis performed by an engineer (responsible to provide the dispositions that concern each discrepancy found in the analyzed part), cost and warranty analysis (based on the aircraft serial or register number it is possible to track a part whether the repair can be covered under warranty or not), etc.

Balestrassi et al. (2009) mentions that there exists difficulties on defining the factors that control the process. Indeed, it is impractical or impossible to attempt to control more than, say, ten factors; many experiments deal with fewer than 5. The log book process studied in this paper seems to have many factors that can be interesting when running an experiment at first time using as much factors as available, respecting time and cost restrictions. Again, the part factors do not go beyond the typical data, as accumulated flight hours or flight cycles, the group that the part is installed in the aircraft, and the disposition adopted by the repair station.

5 The dataset analysis and results

The efforts to always improve the maintenance techniques have attracted the attention of several research workers. Shankar and Sahani (2003) mentions that increase on availability can be achieved using the maintenance techniques and preventive maintenance programs. The implementation of new techniques can lead to increase of availability and consequently reducing the downtime that generates delay and cancellation situations.

Following this practice, Luxhøj (1999) mentions the failure information follows patterns that are cyclic or repeated. The process behind the studied dataset here can be modeled as a function of various input control variables, with the correspondent output variables. The considered noises are all the influences that contribute negatively to the process behavior.

Taking these premises as baseline, the text mining used in conjunction with artificial neural networks (ANN) as approximation models are promising when compared to standard analytical techniques. This problem basically deals with many variables in the process when the flight crew reports an event, the necessary time to analyze and solve the problem, and also the necessary resources to accomplish the task and return the aircraft safely to its airworthiness condition.

5.1 Dataset pre-processing utilizing text mining

In this paper, the dataset under investigation contains data collected from airlines from 2004 to 2008. The total of collected cases amounts to 15,760; which have been grouped by aircraft operator, aircraft serial number, date, failure mode, time of delay, reported problem, and action reported. As previously mentioned, other inherent factors to the data collected that does not appear in the table header, but must be considered are human errors and uncontrolled variables (noises) such as lack of training, forms erroneously filled out, and etc (Table 1).

Table 1 Total of reported events per year

Year	Reported events
2004	194
2005	2,258
2006	2,011
2007	3,156
2008	8,141
Total	15,760

In the dataset, the variables eligible for text mining techniques are the reported problem and the reported action. To start the text mining, some premises were used, such as the minimum and maximum size of a relevant word, minimum size of indexed word, minimum number of vowels, maximum number of consecutive consonants, maximum number of consecutive vowels, maximum number of consecutive identical characters, maximum number of consecutive punctuations, and minimum and maximum percentage of cases where a word appears. Table 2 shows the parameters:

Besides these initial parameters to be used on the dataset text mining, other filters can also be applied such as valid characters used to begin, to write, and to end a word.

For this analysis, initially, elimination of all words that do not pertain to English language, so it can be possible to eliminate from the valid characters list all Latin Characters, for instance the character cedilla like with it was necessary to perform stemming in Portuguese language (Orengo and Huyck 2001). Proceeding in this direction filters all the words containing a cedilla so it will not be considered for effects of text mining indexing purposes. Excluding, for instance, numbers to begin a word will help to eliminate parameters not relevant such as a seat location (e.g. "Seat 12D"). For indexing purposes it is much more intriguing to know that there is a problem with the category seat instead of a punctual problem located in a particular seat. The stop words and synonyms list complement the filtering capabilities of text mining. Table 3 shows the parameters chosen for the valid characters:

Due to the fact that the dataset presents various particular terms from aircraft, it was constructed a specific dictionary table to translate them into complete words. Such dictionary table comprehends some acronyms, synonyms, common misspelled

Table 2 Filter parameters for text mining

Condition	Valid number of characters
Minimum size of word	3
Maximum size of word	25
Minimum number of word vowels	1
Minimum size of word	3
Maximum number of s	5
Maximum number of consecutive vowels	4
Maximum number of consecutive same characters	2
Maximum number of consecutive punctuations	1
Minimum % of cases word occurrence	3%
Maximum % of cases word occurrence	100%

Table 3 Valid characters for words

Position in the word	Valid characters
Characters for words	–abcdefghijklmnopqrstuvwxyz
Characters to begin words	abcdefghijklmnopqrstuvwxyz
Characters to end words	abcdefghijklmnopqrstuvwxyz

forms, and others, so that the text mining module will count even the misspelled forms as correct forms in the final count. Table 4 shows these customized forms:

The result of text mining is presented as a vector representation of problem description and reported action, for each given event. Even though the matrix shows a different number of descriptors, the final matrix takes all the necessary descriptors to describe the whole dataset. Thus, for a determined event where one descriptor is not applied, its count is equal to zero. This is to ensure that the final matrix from text mining has the dimension *mxn*, where *m*means the number of analyzed events and *n* the number of used descriptors. Table 5 shows some of the descriptors, raw frequency, number of documents, and stemmed form just after text mining:

The original dataset has 15,760 as mentioned in Table 1. With the text mining results, it was found the total of 88 word descriptors, which can represent approximately 99.24% of the dataset, i.e. 15,640 failure events (these events had at least one descriptor different than zero). The analysis could be redone, but due to the origin of the dataset it was considered sufficient and the discarded part is negligible. The analysis took approximately 3 min using a Pentium® Centrino Duo 1.6 GHz computer processor complimented with 2 GB of Random Access Memory (Table 6).

Finally, Table 7 shows the final result, where each row is a vector representation of words frequency for each case after the step of text mining. Among the transformations available to represent the occurrences (raw frequency, binary frequency,

Table 4 List of synonyms

Descriptor	Synonyms	Descriptor	Synonyms
Adjust	Readjusted, repositionned, repositioned	Lach	Latch
Aircraft	Acft, airplane	Light	Lt, lamp, lights, ligh, ligt, lite
Attach	Resecure, secure, remove, re-secure, re-secured, secu	Maker	Makr
Attendant	Attd	Message	Msg
Battery	Batt, elpu	Operational	Op
Bend	Bent	Overhead	Ovhd
Cart	Trolley	Passenger	Pax
Category	Cat	Power	Pwrd, pwer
Change	Chqd	Recline	Recl, reqd, requ
Check	Ck, chk, ckd	Seam	Sean
Clear	Clean, clear	Seat	Saet
Coffee	Coff	Service	Srvc, svcd, svcs
Emergency	Emer	Serviceable	Svc
Fixed	Mel, reworked	Stay	Tay
Flush	Flus	Tank	Tanl
Forward	Fwd, fwrd	Valve	Valv
Hinge	Hing	Waste	Wast, wate
Hose	Tube, tubo		

Table 5 Analyzed documents

Document	Size	Number of words	Descriptors
1	6	6	Galley drain clog clear galley drain
2	13	13	Cabin hardwar btw seat shear found hardwar instal repair armrest function normal releas
3	10	10	Passeng door ceil shatter bulb clear glass cabin ceil place
4	6	6	Lach place attach screw handl door
5	55	4	Service spill clear spill
6	95	9	Lavatori door broken lock mechan lubric lock mechan check
7	68	8	Pilot copilot life preserv miss replac life preserv
8	68	8	Seat belt miss replac seat belt power aircraft
9	27	4	Clog water pipe repair
10	112	7	Passeng plug lavatori lavatori repair lavatori equip
11	78	5	Broken tray reconnect tray check
12	79	7	Manual releas miss replac miss jump seat
13	63	5	Galley lavatori sink drain clog
14	92	9	Attach lach galley upper miss place attach lach requir
15	70	3	Captain request place
16	116	8	Player volum inaud servic repair check player attach
17	106	7	Water system confus discuss captain discuss water
18	115	9	Water replac water wast system control drain replac hennessi
19	97	9	Tray broken attach tray onboard problem lavatori smoke detector
...	...	...	...
...	...	...	...
...	...	...	...
15751	54	4	Seat doesnt upright adjust
15752	55	3	Seat mechan adjust
15753	37	2	Seat repair
15754	70	3	Seat cabl reattach
15755	50	3	Seat cabl refit
15756	58	4	Seat without button adjust
15757	73	5	Seat fulli upright posit adjust
15758	48	6	Passeng seat cover cabin servic advis
15759	52	3	Seat adjust check
15760	54	5	Seat upright posit seat adjust

log frequency, and inverse document frequency) the inverse document frequency is the adequate to eliminate indexing problems for common words that do not necessarily are representative of the dataset. Such matrix will be later on paired with the remaining variables, such as ATA chapter, failure codes, etc. For the other independent variables, it is not necessary to do text mining, since they are already either pre-classified or have been assigned numeric values.

5.2 Dataset failure pattern recognition utilizing ANN

In the previous topic, it explained how the dataset was pre-processed utilizing text mining. From the raw dataset collected, the variables *reported problem* and *reported action*, initially stored as non-structured text, were converted to frequencies, and

Table 6 Descriptors, frequency, number of documents, and stemmed form

Descriptor	Frequency	Number of documents	Stemmed form	Descriptor	Frequency	Number of documents	Stemmed form
Actuat	158	150	Actuat	Handset	6	6	Handset
Adapt	5	3	Adapt	Heat	116	94	Heat
Adjust	3285	2886	Adjust	Hydrolock	312	295	Hydrolock
Armrest	590	353	Armrest	Insulation	5	4	Insul
Ashtray	1	1	Ashtray	Interphone	30	23	Interphon
Attach	2513	2326	Attach	Knob	41	27	Knob
Attend	1031	805	Attend	Lach	1382	928	Lach
Backlight	3	3	Backlight	Lavatory	4095	2601	Lavatori
Ballast	68	62	Ballast	Leak	507	383	Leak
Belt	488	326	Belt	Mask	198	133	Mask
Blanket	1	1	Blanket	Mast	98	79	Mast
Broken	1892	1823	Broken	Megaphon	68	48	Megaphon
Bumper	11	9	Bumper	Melted	17	17	Melt
Button	596	510	Button	Missing	937	859	Miss
Cabl	891	812	Cabl	Motor	7	7	Motor
Cartridge	5	5	Cartridg	Mount	58	43	Mount
Cushion	125	99	Cushion	Oxygen	124	94	Oxygen
Dimmer	2	2	Dimmer	Passenger	929	826	Passeng
Discharge	3	2	Discharg	Passenger seat	61	56	Passenger seat
...	...	...	...	...	...	...	...
Door	1350	942	Door	Posit	699	671	Posit
Drain	744	500	Drain	Relamp	1060	1048	Relamp
Electr	24	24	Electr	Replac	2751	2644	Replac
Emergency light	47	36	Emergency light	Requir	811	804	Requir
Escutcheon	29	29	Escutcheon	Reset	666	639	Reset
Extinguish	63	56	Extinguish	Seat	8192	5277	Seat
Faucet	325	217	Faucet	Sensor	346	304	Sensor
Firex	61	39	Firex	Servic	2425	2132	Servic
Flashlight	117	65	Flashlight	Sink	702	489	Sink
Flight attendant	544	457	Flight attendant	System	857	709	System
Flush	979	796	Flush	Tighten	156	153	Tighten
Frozen	1194	1049	Frozen	Tray	3215	1962	Tray
Galley	1768	1447	Galley	Upright	669	653	Upright
Halogen	2	1	Halogen	Wast	823	666	Wast
Halon	9	6	Halon	Water	2803	1645	Water

Table 7 Matrix with words frequency

	Actuat		Adapt	Adjust	Armrest		Seat	Servic		System	Tighten		Wast	Water
1	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
2	0.000	...	0.000	0.000	3.799	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
3	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
4	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
5	0.000	...	0.000	0.000	0.000	...	0.000	2.000	...	0.000	0.000	...	0.000	0.000
6	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
7	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
8	0.000	...	0.000	0.000	0.000	...	1.853	0.000	...	0.000	0.000	...	0.000	0.000
9	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	2.260
10	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
15750	0.000	...	0.000	1.698	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15751	0.000	...	0.000	1.698	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15752	0.000	...	0.000	0.000	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15753	0.000	...	0.000	0.000	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15754	0.000	...	0.000	0.000	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15755	0.000	...	0.000	1.698	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15756	0.000	...	0.000	1.698	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15757	0.000	...	0.000	0.000	0.000	...	1.094	2.000	...	0.000	0.000	...	0.000	0.000
15758	0.000	...	0.000	1.698	0.000	...	1.094	0.000	...	0.000	0.000	...	0.000	0.000
15759	0.000	...	0.000	1.698	0.000	...	1.853	0.000	...	0.000	0.000	...	0.000	0.000
15760	0.000	...	0.000	0.000	0.000	...	0.000	0.000	...	0.000	0.000	...	0.000	0.000

finally paired with the remaining variables. Table 8 then summarizes the new dataset structure to be presented to the ANNs:

To start the neural failure pattern recognition, some pre and post processing, selection, and variable coding were done, like transformation of nominal variables to numbers, and normalization. Besides these precautions, putting focus on a specific type of ANN, choose the learning rate (for example small learning rates give slow learning speed and low error rates, whereas big learning rates give the opposite), number of neurons in the hidden layer, training stop criteria, data subsets (to be used for training, validation, and test), and number of epochs used for training.

Table 8 Dataset structure after text mining

	Description
Input variables	
1–2	Year and month for a given register
3–90	Problem/action reported (one column per descriptor)
91	ATA system chapter
92	ATA sub-chapter
Output variables	
93	Interruption (yes/no)
94	Interruption type: • (*Delay*) • (*Cancellation*) • (*Return from runway*)
95	Failure concordance
96	Assigned failure category

The majority of variables are categorical ones, i.e. they are represented by different states like *yes/no*, *hot/cold,* etc. For this dataset, some variables have more than two states. In this case, a special coding procedure was adopted. For example, the variable *interruption type* has three states: *delay, cancellation,* and *return from runway*. So three columns were created for each state and the variables were coded as *delay (1-0-0); cancellation (0-1-0); return from runway (0-0-1).* This precaution, for input and output variables, makes the neural network learning more effective. Moreover, it was considered the cases where the failure was not found, which

Table 9 Input and output variables

		Description
Input variables		
1–88		Descriptors originated from pre-processing phase utilizing text mining (variables: problem reported/ action reported)
89–92		Aircraft model/type (A, B, C e D)
93–97		ATA system chapter
98–119		ATA sub-chapter
Output variables		
120		Interruption? (yes/no)
121		Interruption = *delay*
122		Interruption = *cancellation*
123		Interruption = *return from runway*
124		Failure concordance
		Description—assigned failure category
125	CM01	Attachment problem
126	CM02	Drain problem
127	CM03	Electrical failure
128	CM04	External induced failure
129	CM05	Failure related to galley equipment—e.g. galley inserts failure
130	CM06	Heater problem
131	CM07	Improper cleaning
132	CM08	Improper handling
133	CM09	Improper lubrication
134	CM10	Improper maintenance
135	CM11	Improver servicing
136	CM12	Latch problems
137	CM13	Lock problems
138	CM14	Mechanical failure
139	CM15	NFF (no fault found)
140	CM16	Seat recline mechanism problem
141	CM17	Routine check
142	CM18	Seat belt problem
143	CM19	Wearout
144	CM20	Wrong troubleshooting

represents a NFF (No Fault Found) situation, due to the relevance of this indicator (Söderholm 2005).

Therefore, after the coding and normalization, the dataset can be represented as shown on Table 9:

As per the Table 9, it can be shown that it was possible to condensate a considerable dataset into a reasonable number of variables (119 input variables and 25 output variables).

As the modeling consists of classification of events, based on the problem and action descriptions, it was chosen the MLP network type. This type of network is very flexible to model this sort of problem.

Other important heuristic considerations were also considered to start the ANN modeling (Haykin 1999), as listed below:

1. Update of backpropagation error rate in stochastic (sequential) or batch. One method presents its advantages when compared to the other one. The batch update is better because it presents to the network the errors in the end, when the stochastic model presents the error in the end of each training epoch;
2. Maximize the information content: utilization of an example set to maximize the training error. For this modeling, it was used five years of historic data, and then it was possible the dataset shuffle among epochs and also a great generalization with the neural network. It can be also mentioned that although a considerable number of cases is presented to the ANN, some of the used registers can be suffered misinterpretation from the person who recorded it, and consequently *add* outliers to the model;
3. Normalization of input and outputs: each variable can be pre-processed so that the variables do not interfere during the modeling process of synaptic connection weights. This rule was applied as shown on Table 9.
4. Training by means of *tips: Tips* can be considered as important examples that can be used to enrich de dataset diversity;
5. Learning rate: ideally all the neurons should *learn* at the same learning rate. But according to LeCun (1993) apud Haykin (1999) it is convenient to use low values for learning rate in the last network layers, whereas the first layers should use high values. This happens because for a given neuron, the learning rate must be inversely proportional to the square root of synaptic connections for the referred neuron.

In relation to the frequencies found during the pre-processing phase (text mining), the normalization process was conducted as a way to facilitate the network training. A binary transformation was used, i.e. if the frequency is greater than zero the respective descriptor received “1”; otherwise it received “0”. Additionally, it was used the classical proportion of 2:1:1 (50% of registers used for training, 25% used for test, and 25% used for validation).

In the beginning, the neural network has all the synaptic connection weights with random values (based on a normal distribution with the mean equal to zero and standard deviation equal to one). Then, each register is presented to the neural network, where the inputs excite each neuron and the respective activation functions.

During the validation phase, the data from input variables are presented to obtain a response. The response is then obtained from the neural network is compared to the “actual” response from the dataset. In case that the response obtained from the

Table 10 ANN results

Neural Network arrangement (MLP) Parameters	119-33-25	119-43-25	119-53-25	119-63-25	119-73-25
Activation function					
Layer 1 (input)	Hyperbolic	Hyperbolic	Hyperbolic	Hyperbolic	Hyperbolic
Layer 2 (hidden)	Hyperbolic	Hyperbolic	Hyperbolic	Hyperbolic	Hyperbolic
Layer 3 (output)	Logistic	Logistic	Logistic	Logistic	Logistic
Error classification function	MSE	MSE	MSE	MSE	MSE
Efficiency					
Training	0.930196	0.934415	0.934543	0.937867	0.924827
Test	0.925319	0.926854	0.934271	0.930179	0.937340
Validation	0.931202	0.933503	0.927621	0.923017	0.931969
Error rate					
Training	0.166373	0.167903	0.167232	0.170275	0.174229
Test	0.173939	0.176193	0.176450	0.180737	0.183264
Validation	0.172378	0.178355	0.179919	0.181283	0.182995
Training algorithm	BP	BP	BP	BP	BP
Trained Epochs	987	985	995	997	926

neural network differs from the actual response from the dataset, an error is then registered to calculate the error rate. Finally, during the training phase it was chosen the backpropagation (BP) algorithm.

The summary for the neural networks are shown on Table 10:

The parameters are not chosen arbitrarily. Firstly, the number of neurons varied in the hidden layer, as a way to verify the influence of this parameter on the ANN. Then, it was concluded that this parameter has no direct influence for the problem. For instance, the first column on the Table 10 shows an arrangement 119-33-25, what means 119 input variables, 33 neurons in hidden layer, and 25 output variables.

For the error calculation between the predicted value and the actual value used for training, the predicted value and actual value were each squared. Although, the methods based on the maximum likelihood are recommended for classification problems, the previous function is also used in classification and regression problems with great robustness.

For the neuron activation functions, the hyperbolic function was selected for the input layers and the logistic function to the output layers. The hyperbolic function was chosen because it represents the interval $(-1; 1)$, the same interval where the dataset variables are normalized. On the other hand, the logistic function was used in the output layer because it works in conjunction with the error function during the verification process. It is important to remember that the usage of quadratic error and the logistic functions, even if it is not a statistic error estimation represents a training rate that is faster, stable, and can yield higher correct classification rates.

For the learning rate, it was considered a small value (0.001). The objective of this value is to achieve a slow "convergence" to the point of minimum on the error surface response. In addition to the learning rate, a momentum value of 0.3 was chosen as a manner to the network presents one satisfactory convergence speed (due to available computational resources). Then, 500 epochs to train each set, totalizing 1000 epochs. According to Haykin (1999), empiric experiments have shown that when the learning rate converges to zero, the momentum should converge to one—this produces more speed to the convergence. The opposite is done to the neural network finds the learning stabilization. Different values can generate undesirable effects such as an oscillation on MSE during the training.

For the stop condition, it was chosen a rate of 0.05, this represents that the neural network stops the training when the error is less than this value. It is important to remember that the backpropagation algorithm for the stop criteria can be or cannot be reached, so the error condition can also be verified through the training within different epochs, i.e., when the error degradation indicates cases of *over-fitting* or *over-learning*. In the event that the desirable error rate is not achieved, the other stop criterion is the number of epochs chosen. In practice, what is observed is a gradual drop in the training error when it begins, followed by an error stabilization when epochs increase.

The increase of neurons in the hidden layer neither increased nor decreased the error rate. It remained constant. Therefore for this problem, the simplest network arrangement can be used. Among the trained networks, the best efficiency per subset and error rate resulted. Lastly, the validation figures show the ratio of cases correctly classified which is then divided by the number of total cases separated from the original dataset for validation purposes. In other words, the cases that were not used during the training phase.

6 Conclusions

The present paper proposes a new approach to the failure pattern recognition in a dataset (with registers of failure events reported and originated in aircraft from the commercial regional segment). The literature has other similar works to extract meaningful information with specialized systems, but most of them deal with quantitative data, what makes easier to model a neural network. Considering the present problem when text files are used, the proposed approach is very promising.

Conceptually the model has demonstrated reliable performance, yielding approximately 93% accuracy. This represents a step forward in this area, because in normal circumstances such analysis is performed manually (hand calculated). This work can be used in practical terms as a tool to help decision making process, but when the model is analyzed in terms of individual variables a considerable amount of error rate is observed. This constraint should be considered in a future research in furthering this technology solidifying its practicality and sustained benefits as a robust maintenance program for the aviation industry and the like. Currently, the tool can be used as an extra decision making tool to help airlines keep their fleets in optimal operational conditions and airworthiness.

Acknowledgement The authors would like to thank CNPq and FAPEMIG for the support they gave this work.

References

Air Transport Association of America—ATA (2009). Spec2000: *Chapter 11—Integrated Data Processing Materials Management*. Washington, DC, USA.

Al-Jumaily, H., Martínez, P., Martínez-Fernández, J. L., & Van der Goot, E. (2011). A real time named entity recognition system for Arabic text mining. *Language Resources and Evaluation*, Published online first, 30 April 2011. doi:10.1007/s10579-011-9146-z.

ANAC—National Civil Aviation Agency (2009). Delay statistics in Brazil. http://www.anac.gov.br/. Accessed 01/10/2009.

Andrea, B., & Franco, T. (2011). Mining Bayesian networks out of ontologies. *Journal of Intelligent Information Systems*. Published online first, 13 June 2011. doi:10.1007/s10844-011-0165-4.

Balestrassi, P. P., Popova, E., Paiva, A. P., & Marangon Lima, J. W. (2009). Design of experiments on neural network's training for nonlinear time series forecasting. *Neurocomputing*, *72*(4–6), 1160–1178.

Batyrshin, I. Z., & Sheremetov, L. B. (2008). Perception-based approach to time series data mining. *Applied Soft Computing, 8*, 1211–1221.

Bineid, M., & Fielding, J. P. (2003). Development of a civil aircraft dispatch reliability prediction methodology. *Aircraft Engineering and Aerospace Technology, 75*, 588–594.

Bishop, C. M. (1995). *Neural networks for pattern recognition* (1a ed., Vol. 1, p. 477). New York: Oxford University Press.

Brants, T. (2000). TnT—a statistical part-of speech tagger. In: *Proceedings of the Sixth Applied Natural Language Processing conference* (ANLP 2000) (pp. 224–231). Seattle, WA.

BTS—US Bureau of Transportation—Statistics about delays. http://www.bts.gov/. Accessed 1 Aug 2009.

Chen, T., & Popova, E. (2002). Maintenance policies with two-dimensional warranty. *Reliability Engineering and System Safety*, *77*, 61–69.

Chiang, D., Keh, H., Huang, H., & Chyr, D. (2008). The Chinese text categorization system with association rule and category priority. *Expert Systems with Applications*, *35*, 102–110.

Dai, H. J., Chang, Y. C., Tsai, R. T. H., & Hsu, W. L. (2010). New challenges for biological text-mining in the next decade. *Journal of Computer Science and Technology, 25*(1), 169–179.

Dey, L., & Haque, S. M. (2009). Opinion mining from noisy text data discovering knowledge. *International Journal on Document Analysis and Recognition*, *12*(3), 205–226.

Dumais, S. (2004). Latent semantic analysis, ARIST. *Review of Information Science and Technology*, *38*(4).

Farrero, J. M. C., Tarrés, L. G., & Losilla, C. B. (2002). Optimization of replacement stocks using a maintenance programme derived from reliability studies of production systems. *Industrial Management and Data Systems*, *102*(4), 188–196.

Fernandez, O., Walmsley, R., & Petty, D. J. (2003). A decision support maintenance system—development and Implementation. *International Journal of Quality and Reliability Management*, *20*(8), 965–979.

Hansson, J., Backlund, F., & Lycke, L. (2003). Managing commitment: Increasing the odds for successful implementation of TQM, TPM or RCM. *International Journal of Quality and Reliability Management*, *20*(9), 993–1008.

Haykin, S. (1999). *Neural networks—a comprehensive foundation* (2nd ed., Vol. 1, p. 842). Ontario: Prentice Hall International.

Huang, J., Zhou, B., Wu, Q., Wang, X., & Jia, Y. (2011). Contextual correlation based thread detection in short text message streams. *Journal of Intelligent Information Systems*. Published online first, 24 May 2011. doi:10.1007/s10844-011-0162-7.

Hung, J. L., & Zhang, K. (2011). Examining mobile learning trends 2003–2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher Education*. Published online first, 10 March 2011. doi:10.1007/s12528-011-9044-9.

Jiménez, A., Molina-Solana, M., Berzal, F., & Fajardo, W. (2011). Mining transposed motifs in music. *Journal of Intelligent Information Systems*, *36*(1), 99–115.

Knotts, R. M. H. (1999). Civil aircraft maintenance and support—fault diagnosis from a business perspective. *Journal of Quality in Maintenance Engineering*, *5*(4), 335–347.

Kumar, U. D. (1999). New trends in aircraft reliability and maintenance measures. *Journal of Quality in Maintenance Engineering*, *5*(4), 287–295.

Kurien, K. C., Sekhon, G. S., & Chawla, O. P. (1993). Analysis of aircraft reliability using Monte Carlo simulation. *International Journal of Quality & Reliability Management*, *10*, 2.

LeCun, Y. (1993). Globally trained handwritten word recognizer using spatial representation, convolutional neural networks, and hidden Markov Models. In: *6th Neural Information Processing Systems (NIPS)* (pp. 937–944). Denver, Colorado, USA.

Leitner, F., & Valencia, A. (2008). A text-mining perspective on the requirements for electronically annotated abstracts. *Federation of European Biochemical Societes Letters*, *582*, 1178–1181.

Li, J., Wang, H. J., Zhang, Z., & Zhao, J. L. (2010). A policy-based process mining framework: Mining business policy texts for discovering process models. *Information Systems and E-Business Management*, *8*(2), 169–188.

Losiewicz, P., Oard, D. W., & Kostoff, R. N. (2000). Textual data mining to support science and technology management. *Journal of Intelligent Information Systems*, *15*(2), 99–119.

Luxhøj, J. T. (1999). Trending of equipment inoperability for commercial aircraft. *Reliability Engineering and System Safety*, *64*(3), 365–381.

Macskassy, S. A. (2011). Contextual linking behavior of bloggers: Leveraging text mining to enable topic-based analysis. *Social Network Analysis and Mining*. Published online first, 25 May 2011. doi:10.1007/s13278-011-0026-8.

Madu, C. N. (2000). Competing through maintenance strategies. *International Journal of Quality & Reliability Management*, *17*(9), 937–948.

Magerman, T., Looy, B. V., & Song, X. (2010). Exploring the feasibility and accuracy of latent semantic analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics*, *82*(2), 289–306.

Manco, G., Masciari, E., & Tagarelli, A. (2008). Mining categories for emails via clustering and pattern discovery. *Journal of Intelligent Information Systems*, *30*(2), 153–181.

Manning, C. D., & Schütze, H. (2002). Foundations of statistical natural language processing (5th printing). Cambridge: MIT.

Orengo, V., & Huyck, C. (2001). A stemming algorithm for the portuguese language. In: *Eighth symposium on string processing and information retrieval* (p. 186). Spire.

Prabowo, R., & Thelwall, M. (2006). A comparison of feature selection methods for an evolving RSS feed corpus. *Journal of Information processing and management*, *42*, 1491–1512.

Prado, H. A., & Ferneda, E. (2008). *Emerging technologies of text mining—techniques and applications* (1st ed., Vol. 1, p. 358). Hershey: Information Science Reference—IGI Global.

Punuru, J., & Chen, J. (2011). Learning non-taxonomical semantic relations from domain texts. *Journal of Intelligent Information Systems*. Published online first, 20 January 2011. doi:10.1007/s10844-011-0149-4.

Rajpathak, D., Chougule, R., & Bandyopadhyay, P. (2011). A domain-specific decision support system for knowledge discovery using association and text mining. *Knowledge and Information Systems*, Published online first, 18 May 2011. doi:10.1007/s10115-011-0409-1.

Sarac, A. (2000). *Daily operational aircraft maintenance routing problem.* PhD Dissertation, University of Buffalo at New York, NY.

Sekhon, G. S., Rajpal, P. S., & Shishodia, K. S. (2006). An artificial neural network for modeling reliability, availability, and maintainability of a repairable system. *Reliability Engineering and System Safety*, *91*(7), 809–819.

Shamsinejadbabki, P., & Saraee, M. (2011). A new unsupervised feature selection method for text clustering based on genetic algorithms. *Journal of Intelligent Information Systems*. Published online first, 27 July 2011. doi:10.1007/s10844-011-0172-5.

Shankar, G., & Sahani, V. (2003). Reliability analysis of a maintenance network with repair and preventive maintenance. *International Journal of Quality & Reliability Management*, *20*(2), 268–280.

Shimazaki, K., & Kushida, T. (2010). A preliminary approach to creating an overview of lactoferrin multi-functionality utilizing a text mining method. *BioMetals*, *23*(3), 453–463.

Söderholm, P. (2005). A system view of the No Fault Found (NFF) phenomenon. *Reliability Engineering and System Safety*, *92*, 1–14.

Song, Y. L., & Chen, S. S. (2009). Text mining biomedical literature for constructing gene regulatory networks Interdisciplinary Sciences. *Computational Life Sciences*, *1*(3), 179–186.

Tan, S. (2008). An improved centroid classifier for text categorization. *Expert Systems with Applications*, *35*, 279–285.

Tremblay, M. C., Berndt, D. J., Luther, S. L., Foulis, P. R., & French, D. D. (2009). Identifying fall-related injuries: Text mining the electronic medical record. *Information Technology and Management*, *10*(4), 253–265.

Tsay, R. (2005). *Analysis of financial time series* (2nd ed.). Wiley-Interscience.

Wang, W. M., Cheung, C. F., Lee, W. B., & Kwok, S. K. (2008). Mining knowledge from natural language texts using fuzzy associated concept mapping. *Information Processing and Management*, *44*, 1707–1719.

Yang, Y., & Pedersen, J. (1997). A comparative study on feature selection in text categorization. In: *Proceedings of ICML-97, 14th international conference on machine learning* (pp. 412–420).

Yang, S., Wu, X., Deng, Z., Zhang, M., & Yang, D. (2002). Relative term-frequency based feature selection for text categorization. In: *Proceedings of the first international conference on machine learning and cybernetics* (pp. 1432–1436). Beijing: IEEE.