

Abstract No. 015-0065

Extracting and analyzing information from a large volume of aircraft repair messages

Mingang Fu ¹, Roberto F. Lu ^{1,2}, Richard L. Storch ¹, and Cliff J. Kirkham ²

^{1.} *Department of Industrial and System Engineering, University of Washington, Seattle, WA*

^{2.} *The Boeing Company, Seattle, WA*

Contact:

Telephone: 919-491-7563

Email: mf23@u.washington.edu

Mailing address: Box 352650

University of Washington

Seattle, WA 98195-2650

POMS 21st Annual Conference

Vancouver, Canada

May 7 to May 10, 2010

Extracting and analyzing information from a large volume of aircraft repair messages

Abstract

In this paper, a systematic method is provided to examine messages between the service provider and customers to investigate possible causes of delays of aircraft repair services. The messages contain valuable information that can be used for scheduling repair services. A procedure is developed to extract key information from a large volume of repair messages. The extracted information is analyzed using binary variable multiple regressions.

The relationship between the nature of damages and the time span of repair services is of interest. A vocabulary of key words is developed to capture the nature of damages from multiple aspects. The study shows that some key words and combinations of key words have significant impact on the time span of repair services. Linear and higher order regression models are built with a purpose of predicting time span of repair services based on the nature of damages stated in messages.

Keywords: Repair service, Delay, Time Span, Nature of damages

1. Introduction

Time is a crucial factor in the commercial aviation industry, aircraft availability or daily utilization has a very significant impact on its profitability (Wensveen 2007), therefore, airlines try hard to minimize scheduled and unscheduled out of service time of their aircrafts. However, repair time of commercial aircrafts fluctuates significantly and appears to be very unpredictable due to various factors, which sometimes causes serious delays to the repair services (All activities including maintenance, repair and overhaul are regarded as repair services in this paper), delayed repair services may have undesirable consequences and huge economic loss to airlines. The statistical results of a study (Al-Aqrabawi et al. 2009) on a major U.S. based repair service provider revealed that based on airlines demanded due dates, 82.58% of repair services of were late in the first quarter of 2009; and 75.96% of repair services were late in all of 2008, which indicates the huge difficulty for it to satisfy customers' needs. The delays of repair services will further cause scheduling problems faced by airlines, and eventually result in a substantial loss of profit and decreased levels of service for airlines (Yan and Yang1996).

Extensive efforts have been made to develop resource management technologies aiming to solve this problem in different ways. Farley (2000) introduced the Integrated Diagnosis System (IDS) which provides probable causes of malfunctions and repair suggestions accordingly, by monitoring all messages sent from the airplane's computers to the ground control. Lampe et al. (2004) explored the use of ubiquitous computing technologies to improve maintenance efficiency as well as the level of usability. Cheung and Lu (2005) looked into this problem from a point of view of allocation of labor resources; they proposed a fuzzy AHP (Analytical

Hierarchy Process) approach-based system to solve the problem of the shortage of experienced and qualified engineers.

Many mathematical models were built to find optimal maintenance and overhaul decision. Zhang and Jardine (1998) proposed to use unit-time cost or the total discounted cost as objectives to evaluate describe system improvement due to the maintenance action of an overhaul, which is different from the virtual age approach (Kijima 1989) by considering a direct reduction on the system's failure rate. Pascual and Ortega (2006) minimized global maintenance cost by optimizing life-cycle duration and intervals between overhauls. Wagner et al. (2003) explored the use of agent technologies to coordinate aircraft service teams and presented a new TAEMS (Task Analysis Environment Modeling and Simulation) coordination algorithm.

This paper intends to identify possible causes of delays of aircraft repair services by answering the following two questions:

1. How to obtain desired information from a large volume of repair message written in free text format?

Labor intensive manual documentation and check procedures are still undermining efficiency of the process (Lampe et al. 2004). One problem it causes is the ineffective information management. Repair messages between the service provider and its customers contain useful information that can be used to support critical operational decisions such as scheduling repair tasks. However, all the messages were written in free text and recorded by time order without translation into some more accessible format. As the volume accumulates, it becomes very difficult to extract the useful information in the messages.

2. Is there significant relationship between the requested information and time span of repair service?

The nature of damages is of interest in this paper. For each repair case, the appearances of key words (pre-defined) in related messages are recorded and counted. Does the appearance a certain key word have an impact on the delay of repair services? If yes, how is the relationship between the appearance and time span of repair (linear or nonlinear, positive or negative)?

2. The process of repair service

The subject of this study is a major U.S. based commercial aircraft repair service provider. It doesn't perform reparation, directly; instead, it provides crucial repair information and technical support to its customers: airlines and Maintenance, Repair, and Overhaul (MRO) companies.

The commercial aircraft MRO industry originated in early 1970's, when airlines committing technicians, who were lying idle with in-house work, started to performing third-party contracted repairs (Small Business & Technology Development Center 2002). In order to enhance efficiency and keep their core competitiveness, more and more airlines chose to outsource their aircraft repair work to MRO companies. Despite the economic decline, prospect of this industry is promising. According to Frost & Sullivan (2009), the MRO market earned revenues of \$16.4 billion in 2008. It is estimated that the MRO global market will grow by US\$4.8 billion from 2007 to 2012 (Hexaware Technologies 2007).

Some airlines choose to repair the aircrafts by themselves, thus, they contact the service provider directly, whereas other operators delegate this task to MRO. Fig. 1 illustrates the "direct contact", the process of repair service that is requested by airlines; Fig. 2 describes the "indirect contact", a

similar process but the repair service is requested through MRO on behalf of the airlines.

Statistics from case study of Zeyad et al. (2009) show that repair processes with direct contact on average have shorter time span than those with indirect contact.

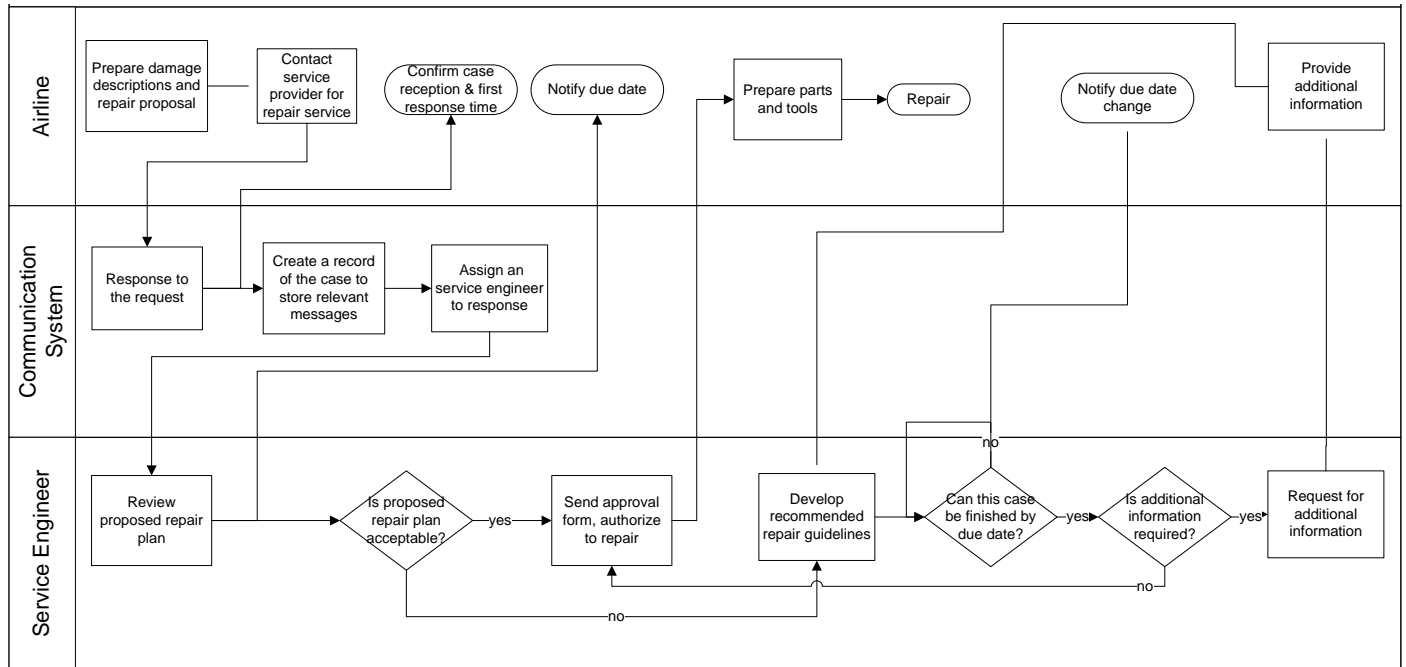


Fig. 1 Direct contact between airline and service provider

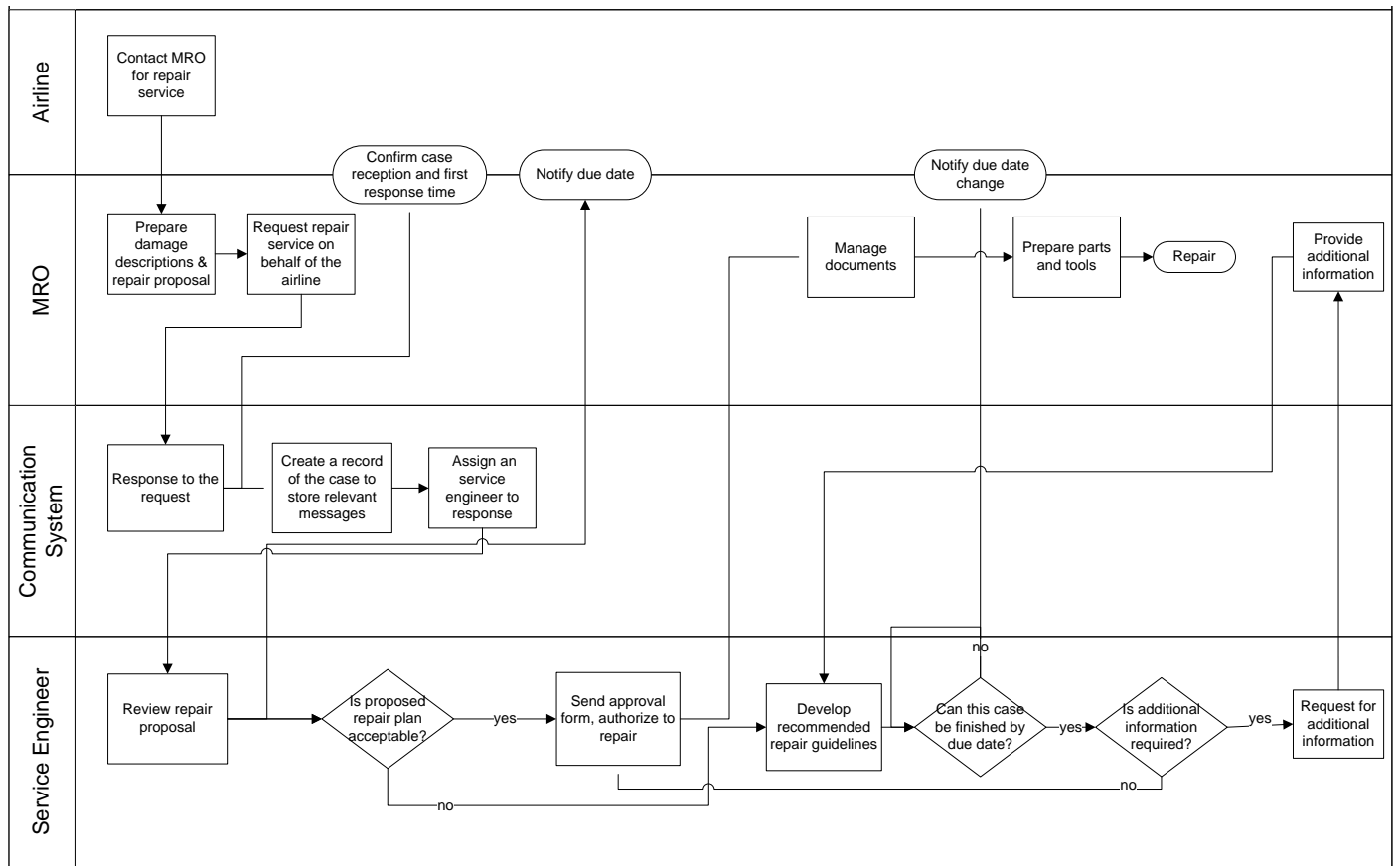


Fig. 2 Indirect contact between airline and service provider

Usually the very first message is sent from the customer to the service provider. It describes the problem, proposes repair plans, and if happened (though not being preferred), the actual repair activities are also included. As soon as the request email is received by communication system, certain information is verified. As long as customer passes verification, the communication system creates a case number to record all related messages, and then sends a customer acknowledgement message, which confirms the receipt and provides a first response date. The messages that follow the acknowledgement email inform the customer that the case has been assigned to a service engineer within the group. The proposed repair plan is reviewed by the service engineer, and if accepted, a signed approval form is sent to authorize the repair; if not,

the customer has to change its repair plan by following repair guidelines provided by the service engineer.

3. The repair messages

This paper analyzes over 1800 repair cases with more than 20,000 messages. The data is stored in a MS Excel file, where each row records an email message with 16 features such as subject, case number, service engineer ID, message body, sent time, and received time, to name a few. Each case involves at least 4 messages, which record the whole process of communication between customer and the service provider for during a case.

Ideally, for a single case, only one final response from the service engineer is desired. However, due to various reasons, many cases are actually delayed for a much longer time than what the service engineers expect. Some are caused by customers who do not provide enough information about the damages and some repair the aircraft without following the correct procedures or do not have the right tools to repair the damages, etc. In some instances, a delay can also be caused by the service provider, who might have provided incorrect repair plans.

Key information about an unexpected delay such as responsible party (customer or service provider), cause, and delay time is contained in the messages. However, it's difficult to extract the desired information from such a large volume of unstructured natural language text data file.

4. Methods

4.1 Information extraction

For each repair service case, the nature of damages, time span and number of messages are of interest. Time span gives measure of lateness of repair service, and the nature of damages is highly suspected to have an impact on time span,

In order to capture the nature of damages, a vocabulary of key words is developed, based on our understanding of the messages and the sample repairs described by Dale et al. (2007). The vocabulary consists of three categories, with a purpose to describe a repair case from three aspects:

- Structure: door, floor, joint, window, etc;
- Position: vertical, horizontal, upper, forward, backward, etc;
- Damage: dent, crack, cutout, etc.

Therefore, extracting information about the nature of damages becomes to be a task to find pre-defined words from unstructured free text. Traditionally, the data to be extracted is typically given by a template which specifies a list of slots to be filled with substrings taken from the document (Mooney and Nahm 2005).

In this paper, the vocabulary is used as a template, which defines a list of key words that are to be filled with substring taken from the message. Fig. 3 shows a shortened message and output of key words extraction. The output includes all the key words that appear in the text.

Example Message

Corrosion was found on the bearstrap common to the fwd entry door cutout skin panel. The corrosion was found on the inner surface of the bearstrap at upper fwd corner of the fwd entry door. The corrosion has been removed by blending the bearstrap on inner surface at 2 locations

per 737-800 SRM 53-00-01. Maximum blend depths 0.045 inch and 0.035 inch. The material removed exceeds SRM allowable damage limits. There is no suitable repair scheme in the SRM for this damage.

Output

Structure: bearstrap, door, panel

Position: upper, fwd, inner

Damage: corrosion, cutout

Fig. 3 Sample message and a template of extracted key words

Every message has an associated time feature, either “sent time” or “received time”, we name it “message time”. Since each repair service case involves multiple messages, the time span of a case is then defined as the one obtained by subtracting the “message time” of the first message from the “message time” of the last messages in messages with the same case number.

The large volume of the repair messages renders infeasibility to perform key words extraction, time span calculation and number of messages count. Since the messages are store in MS Excel format, and MS Excel provides powerful functions, it’s most convenient to perform both tasks in Excel environment.

It takes five steps to finish the three tasks:

- 1 Choose four columns of data from original file: “case number”, “message body”, “sent time”, and “received time”;
- 2 Combine the two columns of “sent time” and “received time” into one column named “message time”;

- 3 Calculate interval between two consequent messages within one case, add intervals together to obtain time span of each case; meanwhile, count number of messages of each case;
- 4 For each case record, keep only the first message; delete the rest of the messages.

This step reduces the data file into a much smaller size, with about 1500 rows.

- 5 Develop a search algorithm to count frequency of every key word in the first message body of every case, mainly by using Excel function “LEN” and “SUBSTITUTE”. Replace all positive frequencies with 1s, and then combine all the key words that appear in one case into a new column named as “key words”.

The above 5 step process is visualized by Fig 4. Output of this process is a Excel data file, in which each row records a repair service case with four features: time span, number of messages, key words, and first message body.

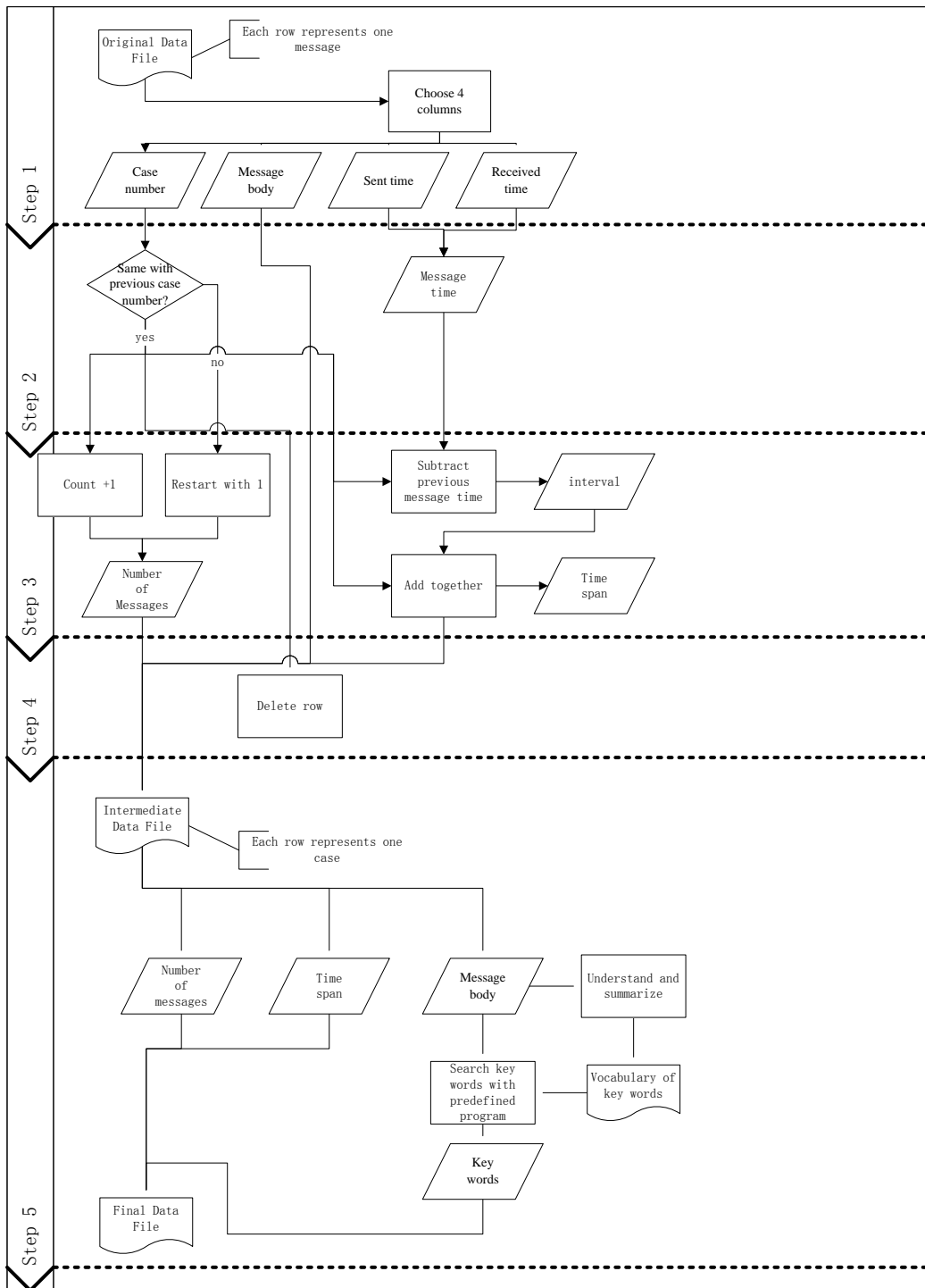


Fig. 4 Process flow of information extraction

4.2 Statistic data analysis

4.2.1 Time span vs. number of messages

Distributions of time span and number of messages are illustrated by Fig. 5 and Fig. 6.

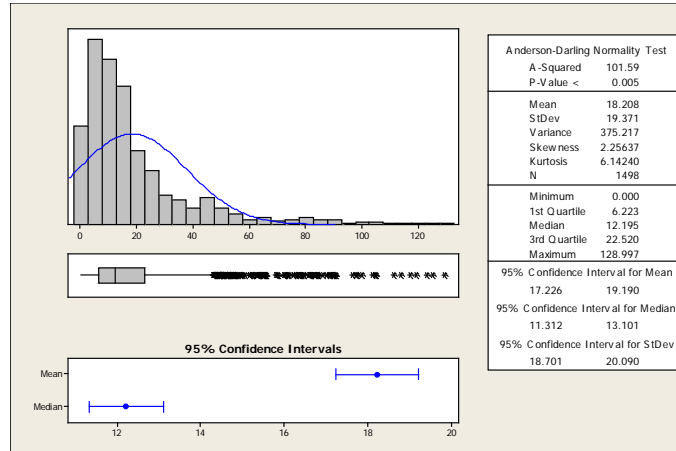


Fig. 5 graphical summary of time span

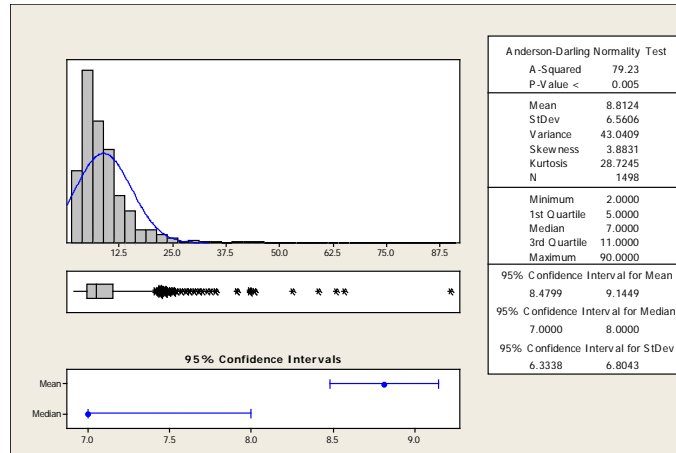


Fig. 6 graphical summary of number of messages

Using a significance level of 0.05, the results of Anderson-Darling normality test (for time span, A-Squared = 101.59, P-Value < 0.005; for number of messages, A-Squared = 79.23, P-Value < 0.005) indicate that both of the time span data and number of messages data follow normal distributions, but only positive parts of the normal distributions. This is probably the result of some abnormal cases, such as abandoned or unfinished (till the date of data collection) cases.

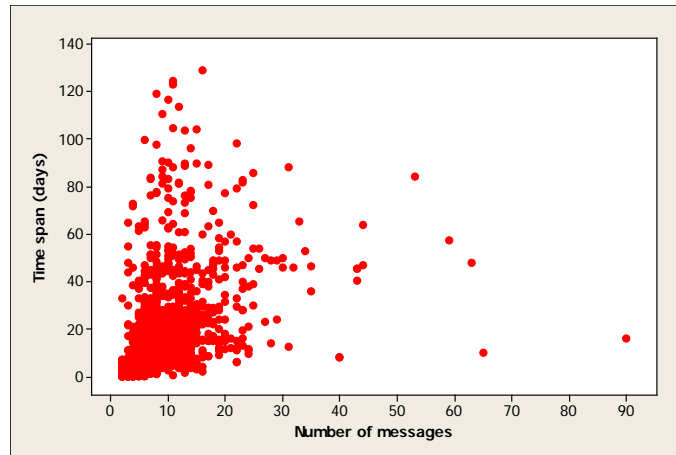


Fig. 7 Scatter plot of time span vs. Number of messages

Correlation between the two variables is of interest. By intuition, there should be positive relation between time span and number of messages. However, the scatter plot between the two variables shows no correlation at all (see Fig. 7).

Considering the abnormal values, Box-Cox transformation is employed to stabilize variance and to improve their correlation (Box and Cox 1964). By applying Box-Cox transformation to both variables, it becomes clear that time span positively correlates to the number of messages (see Fig. 8).

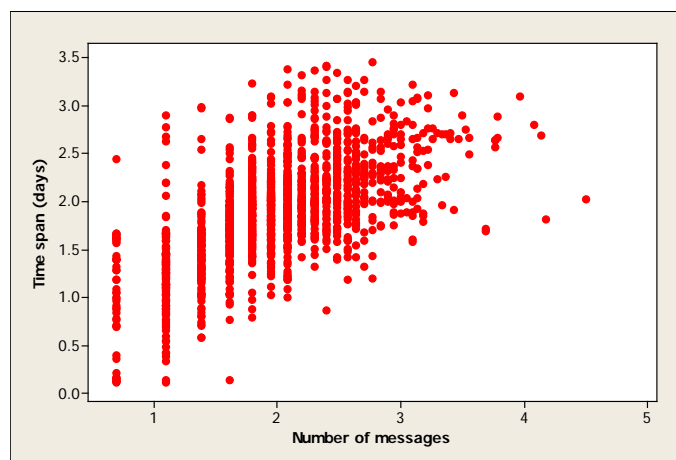


Fig. 8 Scatter plot of time span vs. number of messages after Box-Cox transformation

4.2.2 Average time span vs. frequency

The number of appearances (frequency) of each key word in the vocabulary varies significantly. Out of the total 1500 cases, some key words appear in less than 50 cases (low frequency key word); whereas some appear in over 200 cases (low frequency key words). The impact of key words with different frequencies on time span is of interest. Fig. 9 plots the average time span and frequency for each key word in the vocabulary, each point represents a key word. Average time span of a key word is defined as the average time span of cases that involve word.

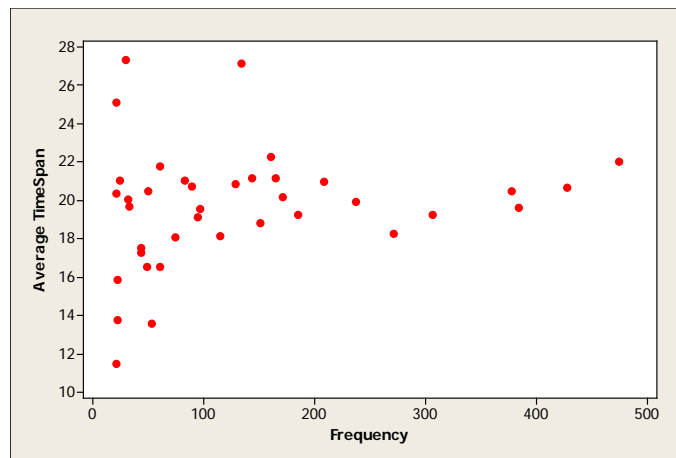


Fig. 9 Scatter plot of Average Time Span vs. Frequency

Some patterns are observed from above scatter plot:

1. High frequency key words have relatively constant average time spans around 20 days;
2. The average time spans of low frequency key words vary widely, thus, need more attention:
 - i. Some have abnormally short (< 14 days) average time spans;
 - ii. Some have abnormally long (> 25 days) average time spans.

To some extent, the patterns are understandable, since high frequency key words usually represent common types of damages whose repair time is relatively stable; whereas low frequency key words represent rare damages, the repair of which have much more uncertainty.

4.2.3 Regression analysis

In this part, regression models are built to test the correlation between the nature of damages and time span (after Box-Cox transformation).

In terms of representation, the feature key words is different from number of messages, it cannot be represented by continuous number. As an alternative, it's treated as categorized variable, specifically, each key word is regarded as a binary variable, a value 1 denotes the appearance of this word, whereas a value 0 denotes no show of this word. Apparently, this problem is a multiple regression with key words as (binary) predictor variables, and time span as a response variable.

A first order regression model is built by Minitab. Using a significance level of 0.05 (which is also used throughout this section), the F test result (3.30, with a p-value of 0.000) of this model indicates that it's statistical insignificant; moreover, its R² value (7.9%) and adjusted R² value (5.5%) show that this model can only explain a small portion of the variation; therefore, it is not suitable to be used for prediction.

To improve this model, the number of variables needs to be reduced, only the variables that have significant impact on the response variable should be kept. There are several ways to reduce the number of key words. One way is to screen out the variables whose p-values of T test (given in regression output) are higher than 0.05; another way is to test significance of all variables one by

one. Since the predictors are discrete binary variables, Analysis of Variance (ANOVA) is used (Aiken et al. 1991). Based on the remaining variables, another first order model is built, the F test result (8.38, with a p-value of 0.000) indicates that this model is statistical significant; however, its R2 value (6.8%) and adjusted R2 value (6.0%) are still too low.

Compared with the first model, F value increases by 5.38, R2 value drops by 1.1%, while adjusted R2 value increases by 0.5%. For individual variables, the p values of T test show significant drop, most of them are lower than 0.05.

Cox (1972) reviewed main methods used in the case of multivariate data with all individual variables are binary. One method proposed by Lancaster (1969) involves a series of chi-squared tests and partitioning of the total chi-squared; another is the second-order method. For the sake of simplicity and effectiveness, second-order method is used. Since there are 13 variables, the number of possible two words combinations out of 13 is $C_{13}^2 = \frac{13!}{2!(13-2)!} = 78$. All possible

combinations among are tested for significance, after abandoning those with p value greater than 0.05, six combinations are left in a second order model, which achieved a R2 value of 7.6% and adjusted R2 value of 6.4%.

Another way is to investigate the correlations between the variables. The correlation matrix shows that some variables have a Pearson Correlation value doesn't give us any hint about the significance of combinations of key words. Some significant combinations have very low correlation value, whereas some highly correlated combinations are not significant at all.

One major disadvantage of the second order model is that it doesn't take account of the interrelationship of three or more variables, which may have impact on response (Cox 1972).

Therefore, a third order regression model which includes three key words combinations becomes desirable. However, it's not feasible to consider all combinations (which in total is

$$C_{13}^3 = \frac{13!}{3!(13-3)!} = 286$$

Direct observation is used to solve this problem. The data file is reduced

to one which only contains the seriously delayed cases (longer than 40 days). There are 140 cases left, and through observation of these cases, four three words combinations are found to appear frequently. After ANOVA, only one is found to be significant. Compared with second order model, the third order model has higher R2 value and adjusted R2 value, but lower F value, which indicates the decrease of significance.

Type of Model	No. of predictor variables	F value	P value	R2	adjusted R2
1st order	38	3.30	0	7.9%	5.5%
1st order	13	8.38	0	6.8%	6.0%
2nd order with interaction	13	6.38	0	7.6%	6.4%
3rd order with interaction	13	6.26	0	7.8%	6.6%

Table 1 - Comparison among regression models

Table 1 summarizes the key parameters of the models, as we can see, compared with the original model, the second and third order models achieve overall higher adjusted R² values but lower R² values with less variables, which is consistent with the common sense that adding more independent variable can increase R² value, but can not necessarily increase adjusted R² value. The introduction of interactions can

help explain more variation of the system, it also gives a better understand of how the nature of damages impact on time span of repair case, however, it reduces the significance of the models.

5. Conclusion

In the repair service of commercial aircrafts, the communication messages between service provider and customer contain key information about the repair process. This paper develops an information extraction procedure to obtain time span, number of messages, and key words of each repair service case. A vocabulary of key words is defined for a purpose of describing the nature of damages; it is used as template for information extraction.

Normality of output data file is examined. Although both time span data and number of messages data follow normal distributed, their correlation is not clear. However, after Box-Cox transformation, the two variables are found to have positive relationship, which could be used to monitor repair jobs dynamically in practice. Close attention may be needed as the number of messages for a certain case exceeds a pre-defined threshold value.

The investigation on key words between frequency and average time span reveals different impacts of the types of damages on repair time, from a perspective of the number of appearances. Cases with common damages are found to have relatively stable time span; whereas the cases with rare damages. This finding indicates that more attention needs to be paid on the repair jobs dealing with rare types of damages.

The regression models reinforce that the nature of damages is found to have significant relation with the delay of repair service cases. The significant two words combinations and three words combination give better interpretation of the nature damages than single key words. However,

the models fail to be tool for predicting time span, due to the low adjusted R2 values. A Bayesian analysis of multivariate binary data may help to get a better model.

To obtain better results on the causes of delays, we suggest the following improvements:

1. Extract and analysis information such as the responsible part for a delay and customer type of a repair case;
2. Deeper understanding of the nature of damages, better vocabulary of key words;
3. A systematic way to find three or more variable combinations.

Reference

Aiken, L. S., S. G. West, R. R. Reno. 1991. *Multiple regression: testing and interpreting interactions*. Sage, Thousand Oaks, California.

Al-Aqrabawi, Z., D. Fleming, A. Maranan, C. J. Kirkham, R. F. Lu, R. L. Storch. 2009. Modeling of Commercial Airplanes Service Request Process Flows. Working paper, University of Washington, Seattle, Washington.

Box, G. E. P., D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society (Series B)*, 26 (2), 211-252.

Cheung, A., D. Lu. 2005. Expert system for aircraft maintenance services industry. *Journal of Quality in Maintenance Engineering*, 15 (4), 348-358.

Cohen, J., P. Cohen, S. West, L. Aiken. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd Edition)*. Lawrence Erlbaum, Mahwah, New Jersey.

Cox, D.R. 1972. The Analysis of Multivariate Binary Data. *Journal of the Royal Statistical Society (Series C)*, 21 (2), 113-120.

- Farley, B. 2001. Extracting information from free-text aircraft repair notes. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 15 (4), 295-305.
- Federal Aviation Administration. 2008. Joint Aircraft System/Component Code Table and Definitions. Oklahoma City, OK. Available at http://av-info.faa.gov/sdrx/documents/JASC_Code.doc
- Frost & Sullivan. 2009. Commercial Aircraft MRO Market Flies into a Headwind of Financial Constraints. PR Newswire, Mountain View, California.
- Hexaware Technologies. 2007. Business Analytics for the Airlines MRO Industry. Jamesburg, New Jersey.
- Johnson, D., R. Lockhart. 2007. Approved Versus Acceptable Repair Data: How to Make Sure You Have What You Need. *aero quarterly*.
- Kijima, M. 1989. Some results for repairable systems with general repair. *Journal of Applied Probability*, 26 (1), 89-102.
- Lancaster, H. O. 1969. Contingency tables of higher dimensions. *Bulletin of the International Statistical Institute*, 43 (1), 143-151.
- Lampe, M., M. Strassner, E. Fleisch. 2004. A Ubiquitous Computing Environment for Aircraft Maintenance. *2004 ACM Symposium on Applied computing*, 1586-1592.
- Mooney, R. J., U. Y. Nahm. 2005. Text Mining with Information Extraction. *Multilingualism and Electronic Language Management : Proceedings of the 4th International MIDP colloquium*, 141-160.

Pascual, R., J. H. Ortega. 2006. Optimal replacement and overhaul decisions with imperfect maintenance and warranty contracts. *Reliability Engineering and System Safety*, 91 (2), 241-248.

Small Business & Technology Development Center. 2002. Analysis of Current and Potential Activity Aircraft Maintenance, Repair, & Overhaul Industry in North Carolina. Business & Research Services, Raleigh, North Carolina.

Wagner, T., V. Guralnik, J. Phelps. 2003. A Key Based Coordination Algorithm for Dynamic Readiness and Repair Service Coordination. *Second international joint conference on Autonomous agents and multiagent systems*, 757-764.

Wensveen, J. G. 2007. *Air transportation: a management perspective*. Ashgate, Hampshire, England.

Yan, S., D. H. Yang. 1996. A decision support framework for handling schedule perturbation. *Transportation Research Part B: Methodological*, 30 (6), 405-419.

Zhang, F., A. K. S. Jardine. 1998. Optimal maintenance models with minimal repair, periodic overhaul and complete renewal. *IIE Transactions*, 30 (12), 1109-1119.