Int J Syst Assur Eng Manag (Apr-June 2013) 4(2):182–192
DOI 10.1007/s13198-013-0148-7

ORIGINAL ARTICLE

A data mining model to identify inefficient maintenance activities

Davood Mosaddar · Amir Abbas Shojaie

Received: 29 September 2012 / Revised: 28 December 2012 / Published online: 20 February 2013

Abstract Nowadays the data collected in the process in maintenance systems comprise a big portion of the related databases. Analyzing these maintenance data provides the firms, enterprises and organizations with a tremendous competitive edge both in manufacturing and service sectors. As maintenance management is a costly and inevitable part of the organization, ensuring that the maintenance activities are performed in an effective manner, is of outmost importance. In other words, organizations can precede with the cost reduction operations, for instance, if and only if the unproductive maintenance activities and processes can be identified. Subsequently, rectifying or removing these kinds of activities or taking other means of modification can help enterprises and organizations to reduce their costs. Data mining is known to be an excellent tool which helps the decision makers to discover the hidden knowledge and patterns when dealing with a large amount of data. Seeing a gap in the related literature reviewed and in order to fill it, this study proposes a data mining based model to identify the unproductive maintenance activities in a maintenance system. By identifying specific inefficient maintenance activities, this model supports the maintenance decision makers to set goals to make amendments in the maintenance systems under their supervisions. Consequently, the organizations can focus on rectifying these fruitless activities and therefore reducing the costs associated with performing them. Finally, the model was used to identify the unproductive activities in a maintenance system comprising of independent components (an urban bus network).

Keywords Maintenance systems · Inefficient maintenance activities · Data mining · Association rules · Clustering algorithm · Association rules algorithm · CRISP-DM

1 Introduction

Today, the competitive business environment urges the companies to cut down on costs in order to maintain a competitive edge and extend it. Different areas of business systems and enterprises have been excavated in order to find a better and more efficient method to deliver the tasks and therefore to cut down on the costs of the whole system.

Maintenance strategies and activities have always been considered to be a costly part of service and manufacturing systems and different paradigms have emerged in the last century to increase their efficiency and cost-effectiveness (Lindley et al. 2002).

Furthermore, knowledge provides power in many manufacturing and service contexts enabling and facilitating the preservation of valuable heritage, new learning, solving intricate problems, creating core competencies and initiating new situations for both individuals and organizations now and in the future (Choudhary et al. 2007).

Hence, using the untapped knowledge in the databases is of prominent importance. This study attempts to address a research gap in the literature of the applications of data mining methods in maintenance.

D. Mosaddar · A. A. Shojaie (✉)
School of Industrial Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran
e-mail: amir@ashojaie.com

2 Data mining and maintenance management

2.1 Data mining

The huge amounts of data in service and manufacturing databases, which contain large numbers of records, with many attributes that need to be simultaneously explored to discover meaningful and useful information and knowledge, make manual analysis impractical. Knowledge discovery in databases (KDD) and data mining (DM) have therefore become extremely important tools in realizing the objective of intelligent and automated data analysis. Data mining is a particular step in the process of KDD, involving the application of specific algorithms for extracting patterns (models) from data. The additional steps in the KDD process, such as data preparation, data cleaning, data selection, incorporation of appropriate prior knowledge and proper interpretation of the results of mining, ensure that useful knowledge is derived from the data (Mitra et al. 2002).

KDD is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. Data mining (DM) is a particular step in this process, involving the application of specific algorithms for extracting patterns (models) from data. The additional steps in the KDD process, such as data preparation, data selection, data cleaning, incorporation of prior knowledge, and proper interpretation of the results of mining ensure that useful knowledge is derived from the data. An important notion of "interestingness" is usually taken as an overall measure of pattern value, combining validity, novelty, usefulness, simplicity and understandability. As a matter of fact, knowledge in this definition is purely user oriented and domain specific and is determined by whatever function and threshold the user chooses. The role of interestingness is to threshold the huge number of discovered patterns and reports only those which may be of some use (Macgarry 2005).

DM is the process of discovering interesting knowledge from large amounts of data stored in databases, data warehouses, or other information repositories (Han and Kamber 2006).

Generally, data mining tasks can be classified into two categories: descriptive and predictive (Han and Kamber 2006). Descriptive tasks can find the hidden patterns, trends and correlations among data variables, while predictive tasks use existing data for future predictions (Larose 2005).

Both descriptive and predictive tasks can be performed through several modeling. The primary data mining models are classification, clustering, association rules, prediction, sequence discovering and visualization (Han and Kamber 2006), (Larose 2005) and (Mitra et al. 2002).

Also, the kind of knowledge to be mined determines the data mining functions to be performed. Possible kinds of knowledge include concept description (characterization and discrimination), association, classification, clustering, and prediction (Han and Kamber 2006).

2.2 Maintenance management

Maintenance strategies embody two primary strategies (Mobely 2004). The traditional Run-to-Failure strategy which is not reliable and in some cases is dangerous and the preventive maintenance (PM) which is a comprehensive plan based on historical data, manufacturers' recommendations and aimed at carrying out the maintenance effectively and preventing failures.

The emergence of preventive and predictive maintenance methods or Total Productive Maintenance (TPM) has all been indicative of the need to improve on the quality and therefore effectiveness of the maintenance activities and to ensure that they are performed the best way possible to decrease the failure of the equipment and its associated costs. However, the application of different maintenance methods and strategies all will undoubtedly end up with a staggering accumulation of the maintenance data, the analysis of which is unavoidable (Choudhary et al. 2009).

Furthermore, as the data collected in the course of time increases in databases, the traditional statistical analyzing methods would be of little help in discovering the hidden patterns in the databases and the need for a new approach seems inevitable (Harding and Neaga 2001).

On the other hand, not all maintenance activities are conducted properly and efficiently. The identification of these unproductive activities is vital from both financial and technical perspectives. These activities yield no desired effects for the system improvement, despite the costs associated with them.

The problem is exasperated in large scaled manufacturing or service systems where the number of maintenance activities is a lot and also the scale on which they are performed is great in size. Finding these inefficient activities through the huge amount of relevant data using the traditional Exploratory Data Analysis (EDA) or statistical methods alone seems to be a formidable task. This makes using another means seem necessary.

2.3 Application of data mining in maintenance and a research gap

Table 1 shows some of the applications of the data mining techniques in maintenance management. As it can be seen from the table, most of the studies have used a particular data mining task to improve an area in the maintenance management.

In regard to a general and comprehensive approach for knowledge discovery, different combinations of methods

Table 1 Studies in the application of data mining in maintenance management

Modeling	Algorithms used	Tasks performed	Ref.
Classification	Neural networks	Prioritizing medical equipment	(Al-Naimal and Al-Timen 2010)
	Decision tree	Suggesting an appropriate preventive maintenance schedule	(Batanov et al. 1993)
Prediction	Neural networks	Predicting dynamically maintenance interval	(Fei et al. 2010)
	SVM	Predicting system reliability and failures	(Ding et al. 2008)
	Naïve bays, decision tree, instance-base	Predicting aircraft components replacement	(Letourneau et al. 1999)
	Bayesian networks	Predicting the failure probability of certain components	(Cai et al. 2008)
Association Rules	Apriori	Discovering association rules to find out the impact of the faults on each other	(Boahui et al. 2011)
	Apriori	Discovering relations between human factors and man-made mistakes	(Zhang and Yang 2006)
	Link analysis	Discovering how and why components fail	(Meseroll et al. 2007)
	Apriori	Using association rules to find the causes of failure	(Young et al. 2010)
Visualization	–	Proposing a new visualization method to analyze fault trends	(Wright et al. 2001)

have also been used. The use of clustering and cluster analysis along with association rules has been adopted to analyze the efficiency of a maintenance system (Makouee et al. 2012). A data mining based algorithm has also been proposed which employs Exploratory Data Analysis (EDA) along with association rules have been used to enhance maintenance management of medical equipment of a hospital (Mokfi et al. 2011).

Taking into consideration the reviewed literature, it can be concluded that none of the studies have focused on the identification of inefficient maintenance activities i.e. the maintenance activities that do not yield a desired outcome for any given reasons.

Seeing this research gap, a data mining base model is proposed here to identify inefficient maintenance activities. This consequently facilitates and supports the maintenance decision makers to set goals to make amendments in the maintenance systems under their supervisions.

Rest of this paper is organized as follows: first the proposed model is presented. Then an empirical case study is presented. In this case study, using the proposed model, it is attempted to evaluate the effectiveness of the maintenance activities of a service system comprising of individual components (buses) the failure of each does not affect the others. The ineffective activities of this maintenance system are then extracted, hence helping the maintenance managers and staff in their decision making process.

3 Data mining model to identify the inefficient maintenance activities

Many academic efforts have long been centered in the attempt to formulate a general framework for DM (Dzeroski and Struyf 2006). The bulk of these efforts are centered in the definition of a language for DM that can be accepted as a standard, in the same way that SQL was accepted as a standard for relational databases (Han et al. 1996), (Meo 1998), (Imielinski and Virmani 1999), (Sarawagi 2000) and (Bota et al. 2004). Most of the efforts in the industrial field concern mainly the definition of processes/methodologies that can guide the implementation of DM applications and CRISP-DM is considered to be the most popular (Azevedo and Santos 2008).

As a result of this, in this study CRISP-DM was employed as the core of the proposed model for the application of data mining. Figure 1 illustrates the schematic steps in CRISP-DM (Larose 2005) and (SPSS 2007).

Figure 2 depicts the proposed model. The following is the description of the proposed algorithm.

3.1 Business and data understanding

As all DM projects, the first phase of this model is business and data understanding. In fact, the role of the domains' experts in this phase is so crucial that the success of many of the DM and KDD projects lay in their hands. Maintenance environment and features of the company or organization under study, the types and natures of the activities, resources and many other issues can only be enlightened by them.

3.2 Data understanding, preparation and selection

Maintenance data warehouses and databases replete with all sorts of data, some of which have not been touched in years and some might totally be redundant (Han and Kamber 2006). A practical maintenance management

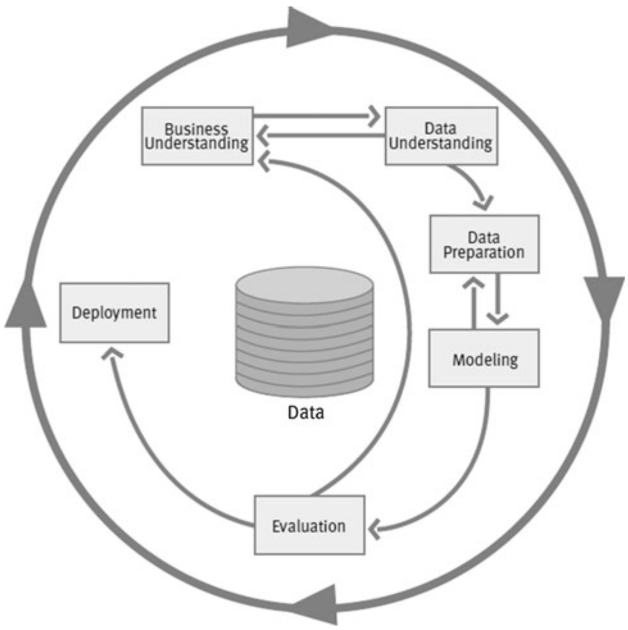


Fig. 1 Phases of the CRISP-DM reference model

system (MMS) may contain many data sources as illustrated in Fig. 3.

Manufacturers' data can be obtained from the users' and servicing manuals; these sources may contain information on lubrication intervals, adjustments and other maintenance activities. Failure data can easily be collected by the maintenance staff (Mokfi et al. 2011). PM data can calculated using the failure data such as the cases with mean time between failure (MTBF) and mean time to failure (MTTF) or by categorizing the maintenance activities according to their nature such as: preventive, corrective, and emergency activities (Lindley et al. 2002).

Maintenance experts may also provide some very useful recommendations on how-to-do of the activities. The data related to the human resources in the maintenance department regarding their skills, experience and training courses, the piece of machinery they have worked on and other related data is another source for data. Some inventory related data such as inventory levels, the prices of the spare parts and lubricants and coolants can also be added to the database.

Another important role played simultaneously by data miner and the domains' experts is data selection phase. The number of fields fed into the model and their significance can enormously change the outcome of the model and the gained results.

Like many databases, maintenance databases may also contain noisy, redundant or missing values (Han and Kamber 2006). Therefore, the next phase in the algorithm is data preprocessing and preparation. Statistical and visualizing methods are used in this stage to determine and identify the outliers, data imputation and data reduction. Then, the prepared data is fed into the data mining phase.

3.3 Modeling

Like all the DM and KDD projects, this phase is in fact the heart of the model. This phase comprises of three different, yet related parts; clustering and association rules along with EDA. Below there is description of each of these:

3.3.1 Clustering

Since in this model, the application of the modeling algorithm must, in essence, partition the related maintenance data into different clusters; in this phase different clustering algorithms can be used.

The choice of the algorithm and the number of clusters depends on the data miner and the domains' experts' opinion.

In cases where the number of clusters cannot be easily determined, methods such as Davies-Bouldin Index can be used (Michaud 1997).

Furthermore, the input variables need to be identified with miners' and domain experts' collaboration.

To analyze the features and properties of the records in the clusters, there are several methods. The most common way is to use a distance function such as the Euclidean Distance. In fact, when the attributes are normalized and are of equal importance, to compute the distance between the two clusters the standard Euclidean Distance is used (Larose 2005) and (Witten et al. 2011).

$$d_{Euclidean}(x,y) = \sqrt{\sum (x_i - y_i)^2}$$

Where $x = x_1, x_2, \ldots, x_m$ and $y = y_1, y_2, \ldots, y_m$ represent the m attribute values of two records.

In addition, Data Exploratory Analysis and other statistical techniques can be used to further find the properties and characteristics of the records in each cluster.

3.3.2 Association rules

Traditionally, association analysis is considered an unsupervised technique, so it has been applied in knowledge discovery modeling. However, studies have shown that knowledge discovery algorithms, such as association rule mining, can also be successfully used for prediction in classification problems too (Li et al. 2001) and (Garci'a et al. 2008).

To find the ineffective activities in the maintenance activities the clusters are placed as the targets of the association rule algorithm and activities as the antecedents. Hence, the activities that have lead the records to be placed

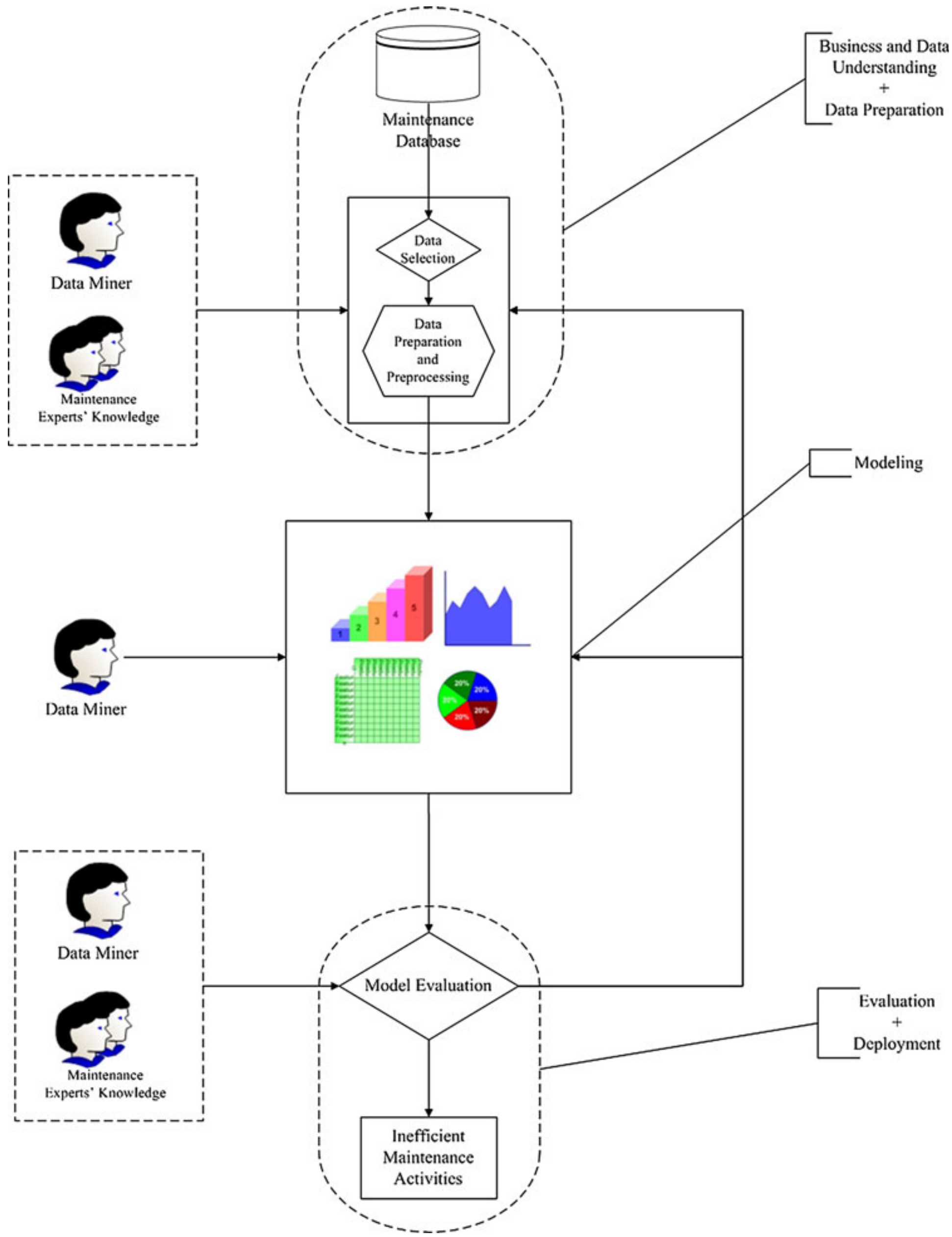


Fig. 2 Data mining based Model to identify the inefficient maintenance activities

in the clusters which poor condition records could be identified.

3.3.3 Exploratory data analysis

Data miners and analysts do not need always have a priori notions of the expected relationships among variables. Especially when confronted with large unknown databases. They often use exploratory data analysis (EDA) or graphical data analysis (Larose 2005).

The use of EDA can greatly assist the miner to further delve into the clusters and find out the properties of the records in them.

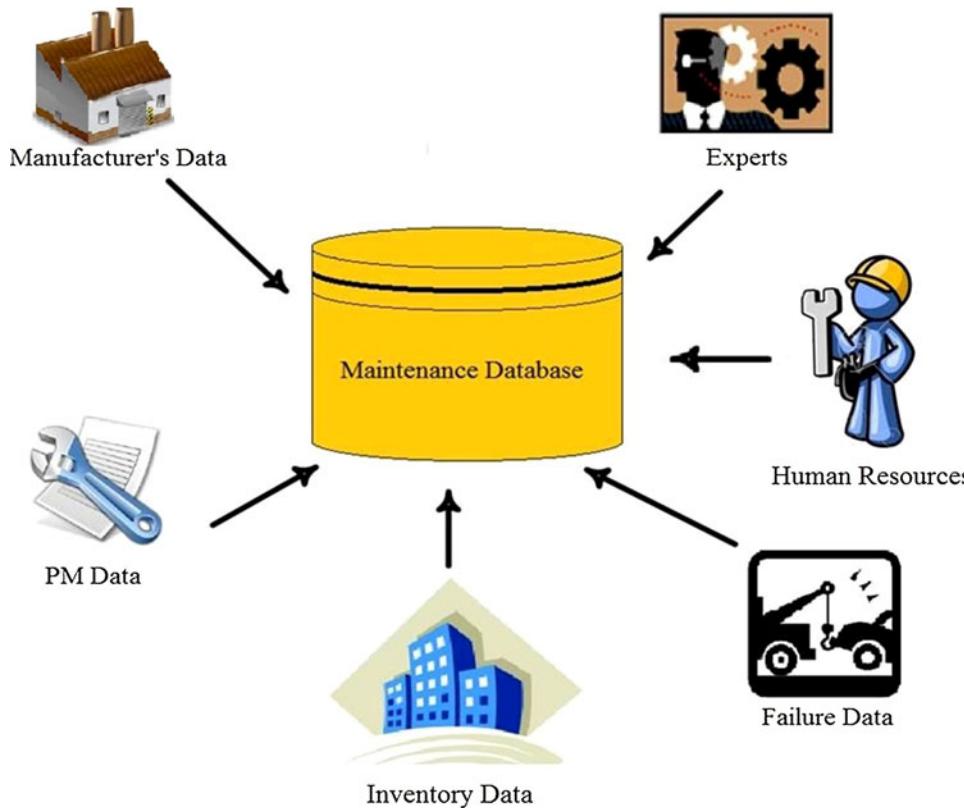
3.4 Evaluation and deployment

The last step in this model is the evaluation of the outcomes and deployment of the findings. Model evaluation is in fact where the outcome of the modeling phase is analyzed in order to see the validity of the outcomes based on their interestingness and usefulness (Macgarry 2005).

Using the domain experts' ideas the validity, applicability and usefulness of the findings can be verified. Figure 4 shows the schematic view of the modeling phase.

As it can be seen in Fig. 2, in the business and data understanding, evaluating and also deployment phase data miners work along with the maintenance experts. Also in

Fig. 3 Maintence data sources



the data preparation phase, because identifying the redundant data in addition to using different techniques to handle the missing data or merging different fields to create new fields to be used in the modeling phase, urges the presence of the domain experts along with data miner.

As shown in Fig. 2, in all the three stages of the model the data miner or data mining team can use the feedbacks taken from each stage to make the needed alterations to get the most desirable results.

Finally, the obtained results are used to support the maintenance management team.

4 The empirical case study

As stated in Sect. 2, in this part an empirical case is presented into further illustrate the proposed algorithm. For this purpose, the maintenance system of the biggest urban bus network in Iran has been chosen.

Below different stages of the algorithm have been used to illustrate the case.

4.1 Business and data understanding

The 196 maintenance activities were performed by 35 personnel in 6 different workshops. The database comprised the collected data for a period of 2 years so that the seasonal behaviors of the records (buses) could all be considered. The urban network consists of 3,000 buses of which 1,100 are operational.

Using the maintenance experts the activities were divided into 3 categories (Lindley et al. 2002):

(1) Emergency or reactive activities: Maintenance activities which take place at the time of a failure in the machinery.
(2) Corrective activities: Maintenance activities which take place when the indications of a failure are detected. They are taken before the occurrence of the actual failure.
(3) Preventive activities: Maintenance activities which are taken to avoid the failures before the detection of the indications.

As it can be seen from the diagram, the percentages of the activities performed in the system are almost equal.

There are 5 models of buses working in the system. Diagram 1 shows the percentages of each model in the system.As it can be seen Model C with over 77 % of the buses is the most common model. Therefore, only this model is chosen for our purpose.

The maintenance activities are performed in 6 different workshops. The percentage of each type of the activities performed in each of the workshops are shown Diagrams 2 and 3.

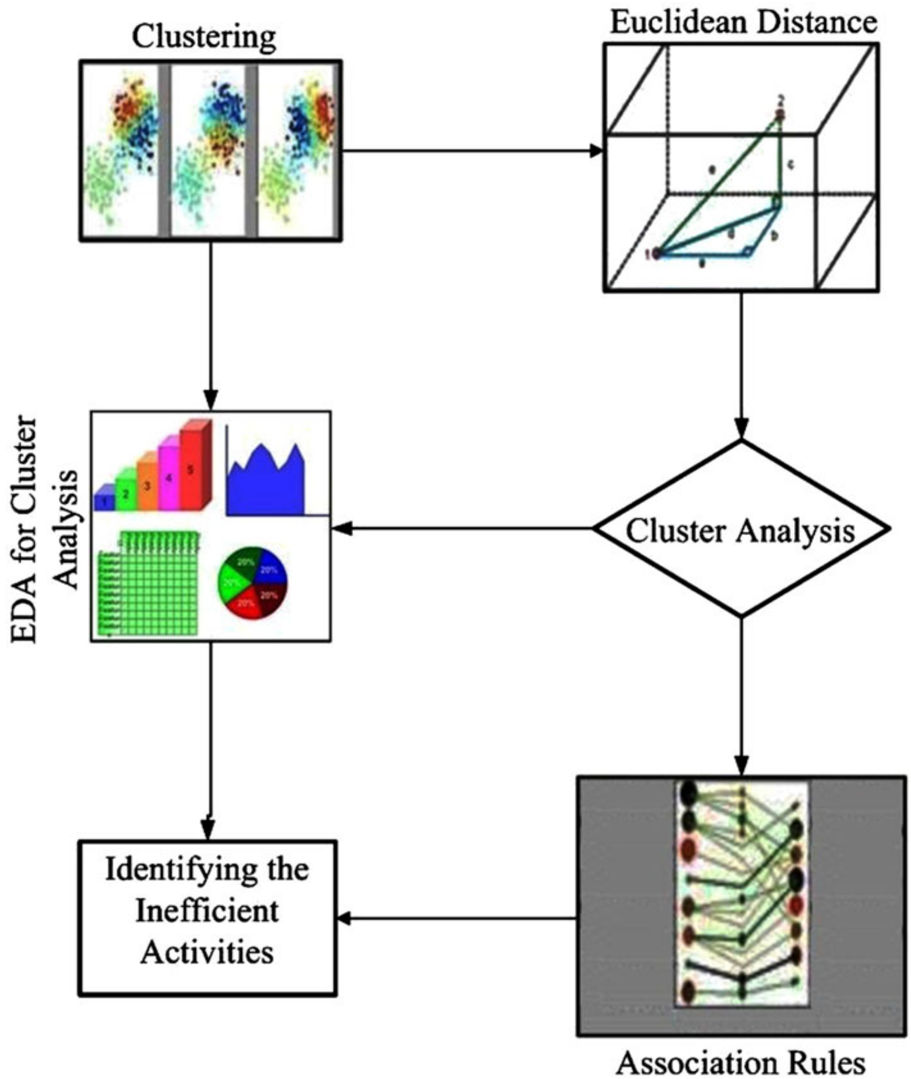


Fig. 4 The modeling phase

4.1.1 Data selection and data preprocessing

As stated earlier this phase mandates total commitment and cooperation of the data miners and the domain experts. Using the previously described database and after removing the redundant, noisy data, the fields that were to be fed into the modeling phase were selected:

1. Sum of preventive activities (which had to be reclassified before being fed into the modeling phase; i.e. because it had to be in same directions as the other input variables.)
2. Sum of corrective activities.
3. Sum of emergency activities.
4. No of references to the workshops over the period of 2 years.
5. Mean Time Between Failure (MTBF) of each record.
6. The average mileage of each record.

4.2 Modeling

4.2.1 Clustering

The choice of clustering algorithm depends on variety of factors but since the maintenance team new about the number of cluster and furthermore because of its efficiency, k-means was selected for implementing this operation (Larose 2005).

Using the k-means algorithm the data was divided into 3 clusters. Diagram 4 shows the number of the records in each cluster.

4.2.2 Cluster comparison

Using the Euclidean Distance the distance between the clusters and an ideal status (0,0,0,0,0,0) were calculated.

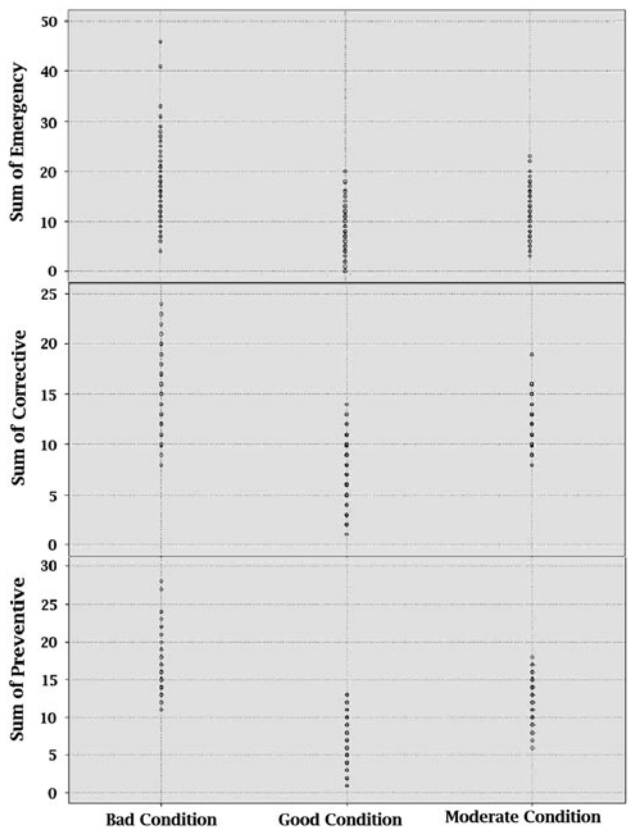


Fig. 5 The 2D scatter plot of sum of activities in the clusters

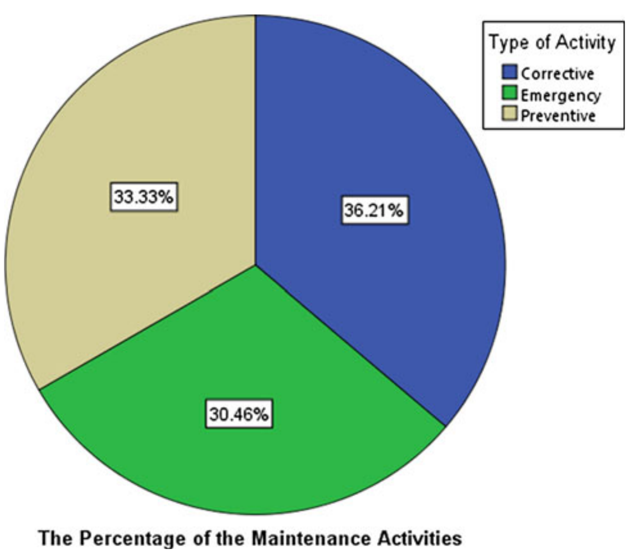


Diagram 1 The percentage of the maintenance activities

The logical conclusion would be that the records in the cluster with the largest distance will have the worst conditions (Makouee et al. 2012).

The clusters were labeled based on the condition of their records to make their recognition easier for the maintenance team. Table 2 gives the description of the clusters.

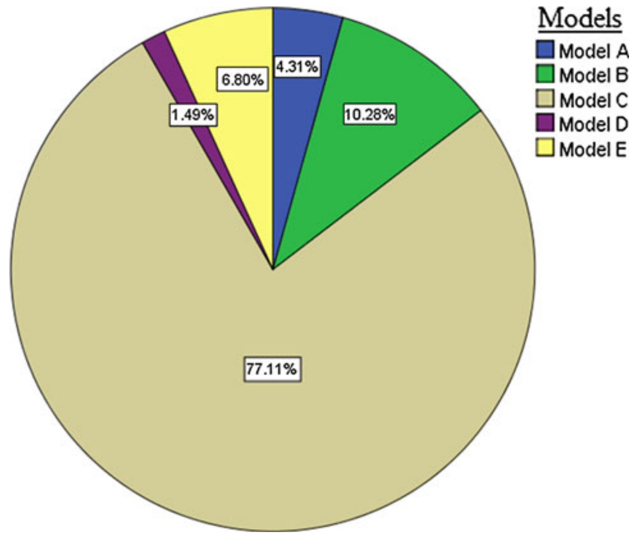


Diagram 2 The percentages of the bus models running in the network

4.2.3 Association rules

As stated above by putting the clusters as the targets (consequents) in an association rule algorithm and maintenance activities as the antecedents, rules can be drawn with the help of which different conditions that have led the records to be placed in each clusters.

The maintenance experts and data miners need to select the rules based on their interestingness the defined thresholds.

There are a number of indices to classify the rules. The first one is the support of the rules which indicates the occurrence probability of antecedents and consequent simultaneously. Confidence of a rule is related to the occurrence probability of consequence contingent upon occurrence of antecedent (Han and Kamber 2006). Generally, the higher these two indices the more reliable is the rule. The other index is lift which is the measure of interestingness of a rule. In other words the higher the lift of a rule, the more interesting it is.

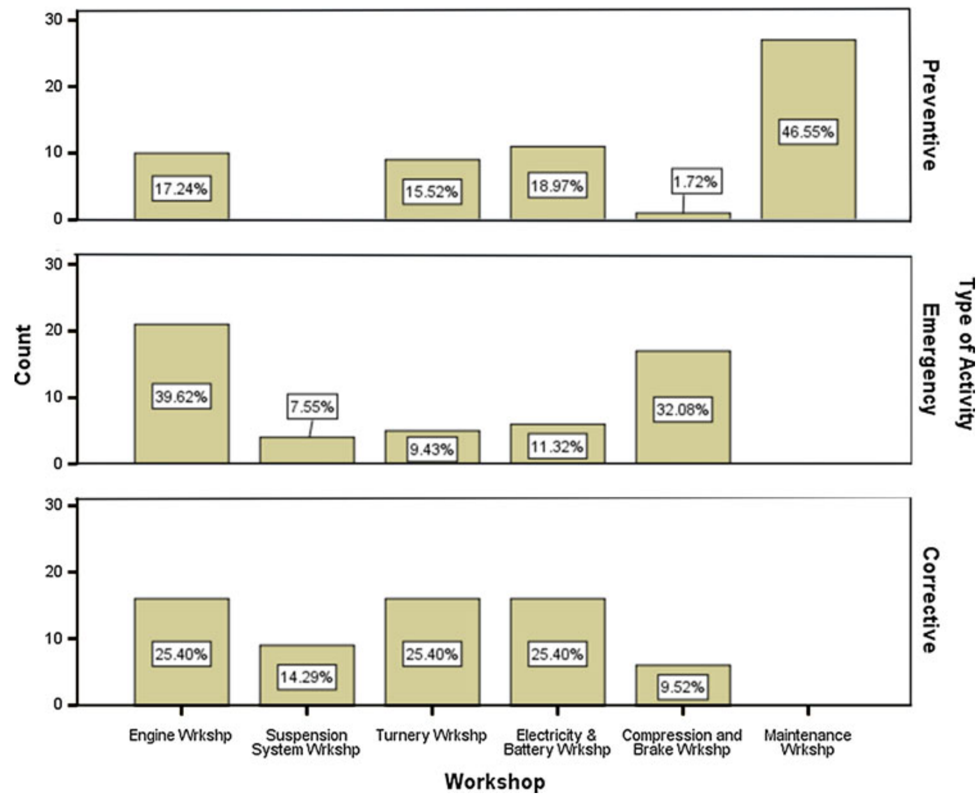
To derive the association rules, the Apriori algorithm was used. Table 3 shows some of the rules extracted from the dataset.

As it can be seen from Table 3, rules 1 and 2 are the ones that have considered the maintenance activities collectively. As rule 1 indicates, when preventive activities on record in a period of 2 years is less than or equal to 150 activities and emergency activities are equal or more than 430, the record is in a bad condition with the confidence of 100 %.

4.2.4 EDA for cluster analysis

In this step of the model, statistical and visualization techniques are used for exploratory data analysis.

Diagram 3 The percentages of each types of activities in each of the workshops



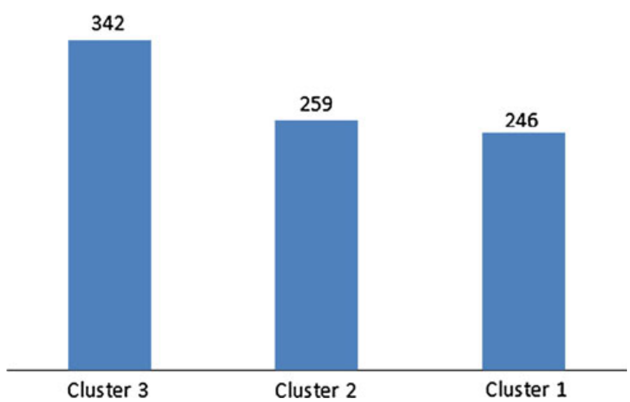


Diagram 4 Number of records in each cluster

Table 2 Euclidean distances of the clusters and their associated labels

Cluster	Euclidean dis.	Labels
1	6.8890	Good condition
2	8.1368	Bad condition
3	7.9927	Moderate conditions

Figure 5 displays the 2D scatter plot which shows the sum of each type of activities performed on the records in each of the clusters. Using these types of plots visually facilitates the understanding of the conditions of the records in clusters. As it can be seen from Fig. 5, the number of activities performed on the records in the cluster Bad Condition is higher than other clusters. Interestingly, the number of preventive maintenance activities that have been performed on this cluster is higher than the other clusters. This might show that the current preventive maintenance approach on the records in this cluster is not effective and the maintenance team is required to further investigate the conditions of the records.

Another test that can be used to see the importance of the activities in relation to the number of failure of each record is Pearson Chi square test.

Pearson is one of the most well-known tests among several chi-tests which tests independence of the target and the predictor without pointing the strength or direction of any existing relationship (SPSS 2007).

Table 4 indicates that theses 11 maintenance activities have strong correlation with the number of failure of the records therefore the maintenance team can make sure that these are done in the most efficient way.

5 Conclusion

Seeing the research gap in the reviewed literature, this study attempted to present a data mining model to identify the inefficient maintenance activities. Seeing the research gap in the existing literature, we tried to propose a model to analyze and identify the inefficient maintenance activities.

Table 3 Examples of the discovered rules

No.	Antecedent	Consequence	Confidence (%)	Support (%)	Lift
1	Preventive activity $\leq$ 150 & Emergency actvty $\geq$ 430	Cluster bad cond.	100	8.11	1.2
2	Corrective activity $\geq$ 150 & Emergency actvty $\geq$ 130	Cluster bad cond.	100	5.12	2.22
3	Activity 102 $\geq$ 5 & activity 32 $\leq$ 15	Cluster good cond.	100	2.41	1.72
4	Activity 130 $\geq$ 14 & activity 182 $\leq$ 50	Cluster good cond.	100	7.15	2.4
5	Actvty 173 $\geq$ 14 & activity 123 $\leq$ 2	Cluster bad cond.	100	6.17	1.23

Table 4 The result of Pearson Chi square test

Rank	Activity	Value	Strength of correlation
1	115	0.89	Strong
2	104	0.816	Strong
3	95	0.815	Strong
4	101	0.785	Strong
5	99	0.709	Strong
6	100	0.708	Strong
7	107	0.700	Strong
8	108	0.618	Strong
9	127	0.564	Strong
10	103	0.471	Strong
11	175	0.455	Strong

Based on the CRISP-DM algorithm, we proposed a model with the use of which the maintenance data could be used to identify the inefficient maintenance activities. In the modeling phase we used clustering to cluster our records (machinery or equipment) into different clusters and using Euclidean Distance we analyzed the conditions of the records in the clusters. By using association rules algorithms we were able to identify the activities that lead the records to be placed in each cluster and finally by using EDA techniques, the inefficient maintenance data could be drawn out.

References

Al-Naimal F, Al-Timen A (2010) A neural network based algorithm for assessing risk priority of medical equipment. 7th International IEEE Multi-Conference on Systems, signals and devices, Amman, 27–30 June 2010

Azevedo A, Santos MF (2008) KDD, SEMMA and CRISP-DM: a parallel overview. IADIS European Conference Data Mining, pp 182–185, Amsterdam, 2008

Batanov D, Nagarur N, Nitikhunkasem P (1993) EXPERT-MM: acknowledge-based system for maintenance management. J Artifi Intell Eng 8:283–291

Boahui J, Yuxin W, Zheng-qing Y (2011) The RESEARCH OF DATA MINING in AHM technology based on association rules. In: Proceedings of IEEE prognostics & system health management, Denver, June 21, 2011

Bota M, Boulicant J-F, Masson C, Meo R (2004) Query languages supporting descriptive rule mining: a comparative study. Database Support for Data Mining Applications. Springer, Heidelberg, p 24

Cai Z, Sun S, Si S, Yannou B (2008) Maintenance management system base on bayesian networks. In: Proceedings of international seminar on information management, 19 Dec 2008

Choudhary AK, Harding JA, Lin HK (2007) Engineering moderator to universal knowledge moderator for moderating lcollaborative. GLOGIFT 07. UP Technical University, Noida, pp 529–537

Choudhary AK, Harding JA, Tiwari MK (2009) Data mining in manufacturing: a review based on the kind of knowledge. J Intell Manuf 20:501–521

Ding F, He Z, Zi Y, Chen X, Tan J, Cao H, et al. (2008). Application of support vector machine for equipment reliability forcasting. International conference on industrial informatics, July 2008

Dzeroski S, Struyf J (2006) Towards a general framework for data mining. Knowl Discov Inductive Databases 4747:259–300

Fei Y, Zhang B, Zhu W, Hu J (2010) Methods of pattern extraction and interval prediction for equipment maintenance. In: Proceeding of international conference on computer and information technology, 9 July 2010

Garcı′a MN, Roma′n IR, Peñalvo FI, Bonilla MT (2008) An association rule mining method for estimating the impact of project management policies on software quality, development time and effort. Expert Syst Appl 34:522–529

Han J, Kamber M (2006) Data mining, concepts and techniques. Elsevier, New York

Han J, Fu Y, Wang W, Koperski K, Zaiane O (1996) DMQL: a data mining query language for relational databases. DMKD-96 (SIGMOD-96 Workshop on KDD), 2. Montreal, June 1996

Harding JA, Neaga EI (2001) Data mining techniques for supporting manufacturing enterprise design. International conference on industrial and production management, Quebec, 20–23 August 2001

Imielinski T, Virmani A (1999) Query language for database mining: MSQL. Data Min Knowl 3:373–408

Larose DT (2005) Discovering knowledge in data, an introduction to data mining. Wiley, New York

Letourneau S, Famili F, Matwin S (1999) Data mining to predict aircraft component replacement. J Intell Syst Appl 6:59–66

Li J, Shen H, Topor R (2001) Mining the Smallest Association Rule Set for Predictions. IEEE international conference on data mining (ICDM'01), Silicon Valley, 29 Nov–2 Dec 2001

Lindley R, Higgins PE, Mobley RK (2002) Maintenance engineering handbook, 6th edn. McGraw-Hill, New York

Macgarry K (2005) A survey of interestingness measures for knowledge discovery. Knowl Eng Rev 20:39–61

Makouee A, Shojaee AA, Mosaddar D (2012) Clustering and association rules in analyzing the efficiency of maintenance system of an urban bus network. Int J Syst Assur Eng Manage 3:175–183

Meo R (1998) An extension to SQL for mining association rules. Data Min Knowl Disc II:195–224

Meseroll RJ, Kirkos CJ, Shanon RA (2007) Data mining navy flight and maintenance data to affect repair. In: Proceeding of IEEE AUTOTESCON, 17–20 Sept 2007

Michaud P (1997) Clustering techniques. Future Gener Comput Syst 13(2):135–147

Mitra S, Pal SK, Mitra P (2002) Data mining in soft computing framework: a survey. IEEE Trans Neural Netw 13(1):3–14

Mokfi T, Almaeenejad M, Sedighi MM (2011) A data mining based algorithm to enhance maintenance management: a medical equipment case study. First international conference on informatics and computational intelligence, pp 74–80. IEEE Computer Society

Sarawagi S (2000) Integrating association rule mining with relational database systems: alternatives and implications. Data Min Knowl Discov IV:89–125

SPSS I (2007) Clementine data mining guide, version 12.0

Witten IH, Frank E, Hall MA (2011) Data mining, practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, Burlington

Wright RG, Kirkland LV, Cicchain J, Deng Y, Dowd N, Hartmuller T, et al. (2001). Maintenance data mining and visualization for fault trend analysis. In: AUTOTESTCON IEEE Proceeding of systems readiness technology conference, 20–23 Aug 2011

Mobely RK (2004) Maintenance fundamentals, 2nd edn. Elsevier, Burlington

Young T, Fehskens M, Pujara P, Burger M, Edwards G. (2010) Utilizing data mining to influence maintenance. In: Proceeding of IEEE AUTOESTCON2010, Orlando, 13–16 Sept 2010

Zhang R, Yang J (2006) Association rules based research on man-made mistakes in aviation maintenance: a case study. In: Proceedings of the sixth international conference on intelligent systems design and applications, Jinan, 2006