Annals of DAAAM for 2010 & Proceedings of the 21st International DAAAM Symposium, Volume 21, No. 1, ISSN 1726-9679
ISBN 978-3-901509-73-5, Editor B. Katalinic, Published by DAAAM International, Vienna, Austria, EU, 2010
Make Harmony Between Technology and Nature, and Your Mind will Fly Free as a Bird
DAAAM Symposium

UTILISATION OF DATA MINING PRINCIPLES IN MAINTENANCE PLANNING

BJORKLUND, S[ergey]

Abstract: *This article is considering data mining techniques for maintenance planning. Author is proposing clustering as a possible data mining method for dividing machines and single components of machines into groups. These groups or clusters are much easier to service if they are divided by the similarity. This article proposes the method of applying dita mining concepts for clustering machines by the cost of service and cost of failure consequences.*
Key words: *maintenance, data mining, clustering, planning, optimisation*

1. INTRODUCTION

Maintenance scheduling utilises raw and unprocessed data about machine failures records for composing the sequence of maintenance activities i.e. maintenance plan. The raw data in this context means number of machines, number of failures, cost of failure and amortisation. The processing of data means the development of maintenance plan. The sequence in which machines should be serviced can be either defined by using the experience of the maintenance personnel or calculated using operational research concepts. In this context the utilisation of data mining techniques is essential for obtaining informative and adequate picture about the current situation with maintenance in the production enterprise.

The studies connected to maintenance usually propose to combine reactive, proactive and planning approach when the servicing of machines is the issue (Sachdeva et al., 2008). Traditional approach is proposing to handle machines servicing correspondingly to seriousness of possible failure. That means the machines with the most critical possible failures are serviced by maintenance plan, machines with moderate seriousness of potential problems are proactively serviced and failures with not so critical consequences are handled in reactive way. The combination of these three concepts together with understanding of organisational core task by maintenance employees provides the opportunity to develop maintenance plan and defines which machines should be serviced and when.

Maintenance plan can be also calculated. One of the possible solutions is to collect information about manufacturing systems and obtain several parameters required, for example, for periodic maintenance problem, preventing maintenance planning model (Aghezzaf et al., 2007) or other models.

This approach has a disadvantage. The information for development and calculation of the maintenance plan is hard to obtain (Abbass et al., 2001). Even in case when the information is obtained, it is difficult to handle it because the data mining techniques are underestimated so far for solving this kind of problems.

2. DATA NEED FOR MAINTENANCE PLANNING

Manufactuing system consists of various machines and machines consist of components, which can brake down if not serviced in the right time. Machines and components in manufacturing system are represented on the Figure 1.

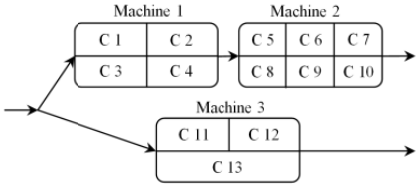

Fig. 1. The part of manufacturing system: machines and components

For developing maintenance plan based on experience and common sense of employees and which is not calculated, the complex data, for example, in the form of spreadsheet, should be collected and analysed. If such data is collected, the maintenance personnel usually try to find such failures, which are the most severe from the point of view of costs. On the manufacturing enterprises this combination of three approaches is usually made on a base of maintenance personnel experience on this production system.

The calculation on maintenance plan requires solving at least one optimisation problem of operations research, which requires formulation of objective function, decision variables and constraints. Maintenance plan in this context is represented by combination of decision variables, objective function is to minimise total cost of maintenance and constraint is the maintenance budget or maximum number of machines which can be serviced during single period. Is there any method which allows to develop maintenance plan more precisely than by maintenance personnel experience, but not that precise than by calculations?

3. DATA MINING TECHNIQUES

Data mining is the process of revealing the certain features from uncategorized set of entities parameters. Using clustering, seriation, searching for assossiation rules and other method it is possible to handle the data in context of gaining information about similar entities in the population, the centroids in clusters, make a decision about to which clusters belongs new entity which entered population. Moreover the variety of data mining techniques utilised for analysing different aspects of business as well as social interrelations is high.

However the opportunities for usage data mining techniques in maintenance planning and scheduling is so far underestimated. These principles allow clustering the components in the production system.

4. CLUSTERING

Clustering is one of the most commonly used methods in data mining practise. Raw data in the tables is usually represented by the entities and parameters. The main task for clustering is to divide the entities in the way that entities in single cluster are possibly similar and entities in two clusters are possible different. For higher dimensional data, a popular measure is the Minkowski metric,

(1)

where is the dimensionality of the data, and are two comparable parameters, and are -values of these parameters. The Euclidean distance is a special case where , while Manhattan metric has (Simovici & Djeraba, 2008). However, there are no general theoretical guidelines for selecting a measure for any given application.

5. ASSESSMENT OF MACHINES IMPORTANCE

Let the be the index of number of components in manufacturing system, – the number of machines and be the number of clusters, which was obtained in the process of clustering. Obviously . This means, the level of importance of component failure can be graded from 1 to . Level of machine importance can be then calculated as sum of importances of all components in this very machine:

(2)

That means every machine gets exact grade which indicate the level of importance. Machines can be then grouped by this level and the maintenance of the groups of machines can be done according to the level of importance and therefore various approaches can be used: reactive, proactive, planning.

6. PRINCIPLES IN PRACTICE

Broken component	Corresponding machine	Cost of component broken	Cost of replacement	Cost of not produced units	Frequency of occurrence per week	Cost of serv. per week
C1	Mach. 1	1	4	90	3	3
C2	Mach. 2	1	5	50	1	10
C3	Mach. 3	3	6	10	0,1	1
...	...	...	...	...	...	...
C	Mach. M	3	6	10	0,1	1

Tab. 1. Raw data of failures track records

The mathematical description of obtained result can be delivered as follows. Table 1 has entities – the possible failures in the production system and the parameters of the entities. In context of clustering it is not interesting to consider neither the strength of dependencies between two parameters of entities nor the network of parameters as it is possible using Bayesian modelling. In this case it is important to divide entities into clusters as entities represent failures.

The next step is the clustering of components and correspondingly the failures. This can be done using the mathematical equations, but there also exists software for handling raw data and clustering it in various ways. There is not so much difference between clustering methods and literature is not giving the proposals about selection of clustering methods (Xu & Wunsch 2009). However author proposes to use Hamming method for clustering.

As every failure corresponds to a single component of the machine, it is possible to cluster failure by the degree of importance. Similarity in clustering can be measured by 19 different methods in PAST software. There also exists software which allows clustering the data into the prescribed number of bins or in other words, in clusters, Figure 2. In this case it is more important to know the tree of clusters. Using the tree the level of precision can be observed.

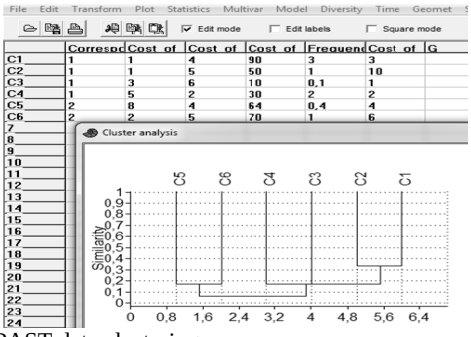

Fig. 2. PAST data clustering

The next step is the assessment of the machine servicing importance.

Corresponding machine	No. of comp. with importance class 1	No. of comp. with importance class 2	...	No. of comp. with importance class N	Sum of importances
Mach. 1	1	1	...	2	7
Mach. 2	1	3	...	1	9
Mach. 3	2	0	...	0	4
...	...	...	...	...	...
Mach. M	0	1	...	2	6

Tab. 2. Machines ranked by importance

After that it is possible to divide machines by the level of servicing importance and therefore decide, using what method the machine should be serviced. For maintenance personnel it is much simplier than make decisions about servicing of machines on the base of failures track records.

7. CONCLUSION

The methods presented in the acticle help maintenance personnel to make robust maintenance plan without utilisation of decision making theory or optimisation using operational research methods. Principles described in the article can become the tool for production enterprises without computer software for planning maintenance activities and which only have track records of failures occurred in the past.

Limitation on the study is its innovativeness and therefore the lack of readiness of maintenance management to use the principles in practise because of the data collecting difficulties. The data about previous machine breakdowns and services is not usually properly stored and therefore utilisation of this method is limited by the quality of raw data collected previously by maintenance personnel.

In future studies author is going to apply the sensitivity theory and assess the precision of results obtained by the practises in this very acticle.

8. REFERENCES

Abbass, H. A.; Sarker, R. A. & Newton, C. S. (2001). *Data Mining: A Heuristic Approach*, Idea Group Publishing, ISBN:1-930708-25-4, Hershey PA.

Aghezzaf, E.H.; Jamali, M.A. & Ait-Kadi, D. (2007). An integrated production and preventive maintenance planning model. *European Journal of Operational Research*, 181, 2, 679–685, ISSN:0377-2217.

Sachdeva, A.; Kumar, D. & Kumar, P. (2008). Planning and optimizing the maintenance of paper production systems in a paper plant. *Computers & Industrial Engineering*, 55, 4, 817–829, ISSN:0360-8352.

Simovici, D. A. & Djeraba, C. (2008) *Mathematical Tools for Data Mining: Set Theory, Partial Orders, Combinatorics*, Springer, ISBN:978-1-84800-200-5, London.

Xu, R. & Wunsch, D. C. (2009). *Clustering*, John Wiley & Sons, Inc., ISBN:978-0-470-27680-8, New Jersey