

MAINTENANCE BEHAVIOUR-BASED PREDICTION SYSTEM USING DATA MINING

Pedro Bastos, Rui Lopes, Luís Pires, Tiago Pedrosa,

Instituto Politécnico de Bragança, Campus de Santa Apolónia, Apartado 1038 - 5301-854 Bragança- Portugal

(bastos@ipb.pt, rlopes@ipb.pt, luica@ipb.pt, pedrosa@ipb.pt,)

Abstract - In the last years we have assisted to several and deep changes in industrial manufacturing. Induced by the need of increasing efficiency, bigger flexibility, better quality and lower costs, it became more complex. The complexity of this new scenario has caused big pressure under enterprises production systems and consequently in its maintenance systems. Manufacturing systems recognize high level costs due equipment breakdown, motivated by the time spent to repair, which corresponds to no production time and scrapyard, and also money spent in repair actions. Usually, enterprises do not share data produced from their maintenance interventions. This investigation intends to create an organizational architecture that integrates data produced in factories on their activities of reactive, predictive and preventive maintenance. The main idea is to develop a decentralized predictive maintenance system based on data mining concepts. Predicting the possibility of breakdowns with bigger accuracy will increase systems reliability.

Keywords - Management, maintenance, data mining, e-collaboration protocols, reference architectures.

I. INTRODUCTION

Global competition has forced manufacturing enterprises to reduce costs, increase quality and efficiency, and be more flexible, in order to sustain their competitiveness [2; 14]. Enterprises have had the need to cope with market expectations, incorporating in their production philosophies new paradigms such as JIT- Just in time, MTO- Make to order, Mass Customization, agile manufacturing or Lean Manufacturing. This posture allows them to compete in markets with a big diversity of products and also big quantities, becoming therefore more competitive. Following this path, enterprises must also incorporate information and communication technologies as one more of their strategic and operational components [10]. Computer networks and telecommunication systems are also a valuable element of the development of virtual collaborations [1] reducing undoubtedly the impact of geography and drawn nations and individuals closer together.

The complexity of this new scenario has caused big pressure under enterprises maintenance systems. Maintenance mission is to make equipment and facilities available when requested. Maintenance function, seen as a non value aggregator one, became more and more requested to contribute to cost reduction, based on bigger and consistent equipment reliability. This perspective is stressed when enterprises existing equipment has an advanced service life. It is expected a profusion of breakdowns at those scenarios and consequently a smaller usability of equipment driving to less productivity.

From an economic perspective, maintenance function is seen to the enterprise as a cost [11]. In fact, experience shows that a major percentage of the overall costs of the business concerns with maintenance [17]. Considering this perspective, decreasing costs with equipment operationalization will increase maintenance productivity and consequently overall productivity [12]. This paper intend to focus on this matter as we believe that globalization could also be an opportunity to increase equipment reliability

II. BACKGROUND

Maintenance

Usually equipment breakdown is directly connected with inspection costs, repair costs as well as costs associated with non production time or equipment non utilization. The quality of the maintenance that is performed affects directly the performance level of a specific business or activity.

Breakdowns, usually, have big impact, because they may affect negatively the outputs of the system, security, environmental health, quality, client service, competitiveness or unit costs. Maintenance policy must be effective in the utilization of resources (people, materials, reposition parts, tools, etc.). Maintenance costs are dependent not only on the maintenance team but also on equipment operators and/or designer.

Technologies used by maintenance function are in an open path of discovery of applications that will stop performance deterioration. Maintenance teams must decide which techniques are more suitable to each situation having in mind that a wise choice must increase systems efficiency as well as reducing global costs. An erroneous choice could be the source of new problems while the existent ones could become worse than they were before.

Maintenance process is usually performed by integration of maintenance and process engineering functions at the phase of selection and application of machines and equipment; and also through proactive actions on those machines and equipments that will necessarily pass by preventive and predictive maintenance [12]. Sometimes changes at the project are also considered. In literature it is possible to find three generic types of maintenance [4; 13]:

- Corrective maintenance, consisting in repair actions when some equipment or machine fails. The equipment is in action until the moment that it fails. At that moment it will be repaired or changed. The main disadvantages of this approach included fluctuant and unpredictable production, high levels of non-conform products and scrapyard as well as high levels of maintenance due interventions motivated by catastrophic fails [16];
- Preventive maintenance, characterized for periodic maintenance operations in order to avoid equipment fails or machinery breakdowns, bases its functionality in manufactures manuals and some heuristics [6];
- Predictive maintenance uses some indicator values to “feel” when some breakdown is eminent. This type of maintenance intends to make interventions on machinery before malicious events may occur.

Data Mining

Nowadays the databases are associated with all the activity areas, resulting in an accumulation of large amounts of data. However, the existence of large quantities of data by itself does not imply the possession of knowledge; this existence is just the first step in obtaining information and knowledge. After data processing the results lead to knowledge that can be used in monitoring, analysis and process prediction [15]. The path which starts with the existence of data to obtain knowledge passes through intermediate steps such as obtaining

information or facts, or by the establishment of rules or values.

Data mining is defined to be the exploration and analysis, by automatic or semiautomatic means, of large quantities of data stored either in databases, data warehouses, or other information repositories to discover interesting knowledge including meaningful patterns and rules [3]. In general, the Data Mining term is the process of analyzing data from different perspectives, categorizing them and summarizing identified relations in order to return useful and usable information. Technically it can be seen as the process to find or identify correlations or patterns among dozens of fields in large relational databases [8], which can also be data warehouses or other information repositories [3]. The most used techniques in Data Mining are:

Artificial Neural Networks: Consisting in a non linear predictive model that learns through training; their structure is similar to a biological neural network;

Decision trees: Structures in the shape of a tree representing sets of decisions. These decisions create rules for classification of data sets. The specific methods of the decision trees include Classification Trees and Regression (CART Classification and Regression Trees) and Chi Square Interactive and Automatic Detection (CHAID Automatic Interaction Detection Chi Square);

Genetic algorithms: Optimization technique that uses processes such as genetic combination, mutation and natural selection, based on the evolution concept;

Neighbourhood closer Method: Technique based on the classification of each record of all data on a class combination of k records that resembles more a history of a data deposit. It is sometimes called technique of k closest neighbours;

Induction Rule: extracts if-then rules from data based on statistical significance.

III. SYSTEM FUNCIONALITY

In literature there are some approaches that use data mining concepts to improve manufacturing activities [7]. It isn't so common to find approaches that use it to improve the capacity of predicting behaviours based on historical data. In fact, if the case is the possibility of a distributed collaboration of independent enterprises sharing data between them, even if they are competitors, the examples are even reduced.

We have found some examples which work on historical failure data and provide suggestions for an appropriate preventive maintenance schedule [7]. At

the same work exist other references to the utilization of data mining techniques, including decision trees, rough sets, regression, and neural networks to predict component failure based on the data collected from the sensors of an aircraft. Other references to neural networks utilization for prediction effects in maintenance scope are present at [9], or references of equipment maintenance using optimized support vector machine existent at.

Databases utilization with statistics is a tool well established in engineering scope. Maintenance activity, as well as other issues, have been influenced and evolved based on information and communication (ICT) developments. Technology evolution has lead to an increasing number of data flowing from all the processes of the organization such as product and process design, material planning and control, assembly, scheduling, maintenance, recycling, etc [7]. On their activity of reactive, preventive and predictive maintenance enterprises produce huge amounts of information that are stored to build historic profiles that can be helpful in future interventions. Data storage is usually made in a local perspective either in electronic or paper support. Nevertheless, it is known that the integration of those tools not always is effectively achieved.

Usually, enterprises do not share data produced from their maintenance interventions. This investigation intends to create an organizational architecture that makes the integration of data produced in factories on their maintenance activities. The main idea is to develop a decentralized predictive maintenance system based on data mining concepts similar to those that support virtual enterprises functionality.

Using IDEF0 method the overall functionality will be deeply explained. IDEF0 method is used to specify function models, representing high-level activities of a process and its decomposition in lower hierarchical sub activities or processes. IDEF0 models portrait a view of the process in terms of Inputs, Controls over the process, Outputs, and Mechanisms acting on the process, named ICOM's. IDEF0 Function Modelling is designed to model decisions, actions, and activities of an organization or system. IDEF0 model notation uses functions and activities abstracted from temporal sequence. The diagrams in this notation show activation (of functions, processes), not flow. More information may be found at [5].

Figure 1 shows the main system functionality in IDEF0 format. The overall system is based on three main processes. Remote data management and

communication (A1); Knowledge prediction system (A2) and Information synthesis and event generation (A3).

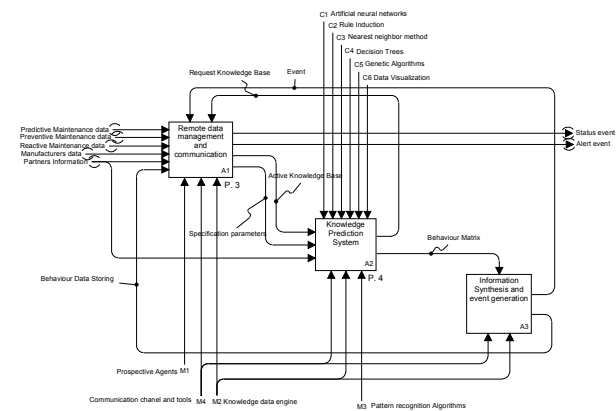


Fig 1: Main System in IDEF0 Format

A1 activity will be responsible for collecting data that will flow from a local perspective to a higher layer. As outputs, this activity will produce an active knowledge base as a control request knowledge base is triggered from A3 activity, and also a set of specification parameters that were used for data preparation. This activity uses four sub-activities (see Figure 2). A11 – Data request analysis and specification. This activity is responsible for analyse and transform the request settings and perform as output a set of specifications that will control activity A12 responsible for data management and selection.

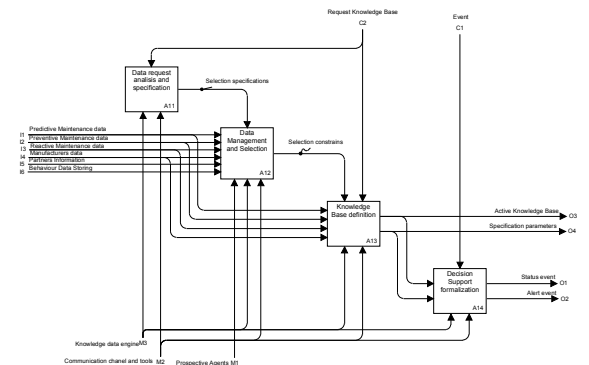


Fig 2: A1 sub-activities in IDEF0 Format

The data collecting process on A12 has to be performed by agents, which will be responsible for adapting and transforming the information. Even when data on the factory floor is collected through maintenance operators (using a formalized internal registry, for example), this information has to be interpreted by software agents, to adapt and transform the data structures into a semantically viable knowledge base. Agents constitute the local

layer, and they have to provide information in a common data structure to the upper layer – the prediction layer. XML is an adequate technology, since it provides a language that can describe the semantics, the syntax as well as the data. Agents, collaborating through an XML language, are responsible for integrating all the information in a common format in a common database (DB). Consequently, the XML language has to transport an appropriated specification language that allows describing the sources structure, the structure of the central DB and the transference schema. The information resulting from the local layer (agents), after being adapted and stored in a common database is ready for the data mining in the prediction layer. From this activity will be outputted the constrains to regulate the production of the active knowledge base, i.e., the knowledge base that will be used to produce knowledge in A13 activity, based on a specific set of constrains. Outputs of A13 activity will affect A2 activity and also A14 activity. A14 activity, named Decision support formalization, will output O1 – Status event and also O2 – Alert event. Those outputs should be used by system partners on theirs preventive maintenance systems as the remote decision support system data.

Both O1 and O2 are triggered by C1, which is the main result of A1 activity recurring to its partner's communication capability.

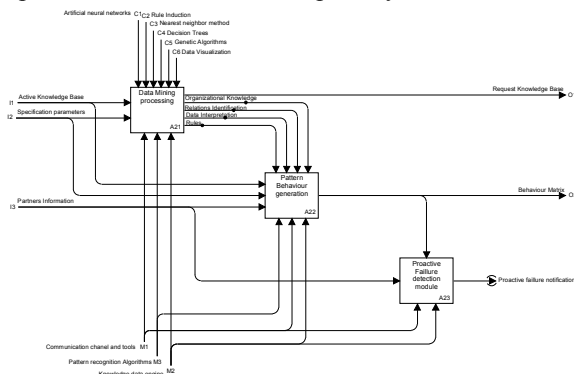


Fig 3: A2 sub-activities in IDEF0 Format

The main module to produce knowledge and infer partner's equipment behaviour is A2, the knowledge prediction system. It is based on three activities. A21 – Data mining processing will use the active knowledge base and the parameters that configure the request as information controlled by data mining controlling methods. The outputs of this activity are organizational knowledge, relations identification, data interpretation and rules. All this information will control the pattern behaviour generation module that will transform the active

knowledge base, specification parameters and partners information in the behaviour matrix. This behaviour matrix will consequently control the proactive failure detection module to transform partner's information in proactive failure notifications. This output intends to advise maintenance responsible to act in their equipment before they fail down. Its functionality is based on deviation analysis of equipment behaviour using data mining algorithms and pattern recognition algorithms.

The new millennium has brought the consolidation of information and communication technologies (ICT) in people life. ICT advances have had deep influence in the way that people live, learn, work or use their spare time. From this reality it is impossible to keep out the concept of trust on those technologies. Also in the proposed system ICT represents the element that incorporates trust. Managing security tasks such as authentication, authorization and access control is a crucial element in environments composed of several and heterogeneous partners. All the members that cooperate on the system must have mutual trust and a blind confidence on the infrastructure. The resultant information is critical. Therefore should be ensuring that the information which flows from the agents to the database is a secure and trustful one.

It is also important that the database server only accept information that come from authorized agents. This validation is very important for stopping a malicious agent/user from flooding the database with erroneous information that will make the prediction model untruthful.

To cope with those needs, the communication should ensure privacy, integrity and non-repudiation. A Public Key Infrastructure can be use to create an environment where the previous problems can be overcome. Therefore, it was decided the utilization of this approach as follows:

- The server will create it's public/private key and act as the Certificate Authority on the system, which means that the server should sign the public key of the authorized agents on the system;
- Each agent has their own private/public key;
- The agents uses theirs private and public server key to encrypt the information;
- The server will only accept data sent by authorized agents;
- The server decrypts using its private key and only after a previous verification.

Based on above specifications the system implements privacy, integrity and non-repudiation.

Privacy is achieved by sending the information encrypted in a way that only the receiver can decrypt. If the information is tampered during the transmission the signature check will fail and notice, assuring that integrity is accomplished. The non-repudiation is achieved once information sent by the agents is signed.

IV. CONCLUSIONS

A globalized world is a place with several challenges and also fulfilled by huge opportunities. In this paper it was presented an architecture that explores the benefits of this globalization and uses new technological means to incorporate more accuracy in a prediction system used by industrial maintenance teams. We believe that the existence of large amounts of data could contribute to find new connections and rules of machines behaviour. Equipment maintenance prediction is an important and widely studied topic since it has a significant impact on equipment maintenance prediction and reliability. The existent huge amount of data from maintenance actions are not fully used for increase the efficiency of maintenance prediction. Data mining seems to be the step forward that will change the actual state.

The proposed system will help enterprises to collect, extract and create knowledge in a way that enterprises will predict with more accuracy the moment to realize maintenance actions and thus improve the productivity of manufacturing process. The innovative point of this system is the capability of collecting and treats data dispersed in different facilities that result from maintenance interventions in different environments. The contribution of machinery manufacturers it is also a relevant contribution.

REFERENCES

- [1] Afsarmanesh, H. and L. Camarinha-Matos (1997). Federated Information Management for Cooperative Virtual Organizations. Proceedings of the 8th International Conference on "Database and Expert System Applications", DEXA'97 (LNCS)1308, Toulouse, France, Springer Verlag.
- [2] Bansal, D., D. Evans and B. Jones (2005). "Application of a real time predictive maintenance system to a production machine system." *International Journal of Machine tools and manufacture*: 1210-1221.
- [3] Chien, C. F., W. C. Wang and J. C. Cheng (2007). Data mining for yield enhancement in semiconductor. *Expert Systems with Applications*: 192-198.
- [4] Chu, C., J. M. Proth and P. Wolff (1998). "Predictive maintenance: The one-unit replacement model." *Int. J. Production Economics*: 285-295.
- [5] Feldmann, C. and J. V. Tieso (1998). *The Practical Guide to Business Process Reengineering Using Idef0*. New York, Dorset House.
- [6] Gharbi, A. and J. P. Kenné (2005). "Maintenance scheduling and production control of multi machine manufacturing systems." *Computers and Industrial engineering*: 693-707.
- [7] Harding, J., M. Shahbas, S. Srinivas, et al. (2006). "Data Mining in manufacturing: A review." *Journal of Manufacturing Science and Engineering* 969-976.
- [8] Lau, H., K. Chinb, K. Punb, et al. (2000). "Decision supporting functionality in a virtual enterprise network." *Expert Systems with Applications* 261-270.
- [9] Luxhoj, J., T. Williams and H. J. Shyur (1997). "Comparison of regression and neural network models for prediction of inspection profiles for aging aircraft." *IIE Transactions* 91-101.
- [10] Moodley, S. (2003). "E-commerce and export markets: Small furniture producers in South Africa." *Journal of Small Business Management* 317-324.
- [11] Narayan, V. (1998). "The raison d'être of maintenance." *Journal of Quality in Maintenance Engineering* 1355-2511.
- [12] Palmer, R. (1999). *Maintenance planning and scheduling handbook*. New Jersey, McGraw Hill.
- [13] Pinjala, S. K., L. Pintelona and A. Vereecke (2006). "An empirical investigation on the relationship between business and maintenance strategies." *Int. J. Production Economics*: 214-229.
- [14] Shiue, Y. R. (2008). "Data-mining-based dynamic dispatching rule selection mechanism for shop floor control systems using a support vector machine approach." *International Journal of Production Research*: 1-22.
- [15] Sikora, R. and S. Piramuthu (2005). "Efficient Genetic Algorithm Based Data Mining Using Feature Selection with Hausdorff Distance." *Information Technology and Management*: 315-331.
- [16] Swanson, L. (2001). "Linking maintenance strategies to performance." *Int. J. Production Economics*: 237-244.
- [17] Wang, F. K. and W. Lee (2001). "Learning curve analysis in total productive maintenance." *The International Journal of Management Science* 29: 491-499.