THE DATA WAREHOUSE SUGGESTION FOR PRODUCTION SYSTEM

VAZAN, P[avel]; KEBISEK, M[ichal]; TANUSKA, P[avol] & JUROVATA, D[ominika]

Abstract: The paper presents the utilization of knowledge discovery process from manufacturing system control by using simulation models. The simulation models of manufacturing systems have been developed to obtain the necessary data about production. Obtained data have to be stored and preprocessed. A data warehouse solution is suitable for this purpose. The various analyses can be performed over stored and preprocessed data by using the process of knowledge discovering in databases. The aim of this process is to obtain new, valid, comprehensible and potentially useful knowledge from the production system
Key words: knowledge discovery in databases, data mining, simulation model, production system, data warehouse

1. INTRODUCTION

The explosive increasing possibilities of generating and storing data is typical for the end of 20th and beginning of 21st century. A lot of organizations cumulate data into their databases, and what is truly necessary, is information. The vision about an achievement of information companies, alternatively least utilization of strategic strength deposited in data sources, requests not only new tools and new methods, but mainly new style of thinking. However the problem isn't only in the elaborating of new models. In a matter of fact idea about that, look like neither obtain information about objects, their behavior, requirements, secretion dependences and moreover (Fayyad et al., 1996).

An analysis and market control, business analyses, risk management and discovering fakes that are additional examples fields of operative activities that need an ideal accomplishment for staff which makes these activities (Trnka, 2010). A control of production process is an activity which needs a perfect knowledge of environment and all activities, which are related with production process. It is called the controlling of knowledge or knowledge control system. Responsible people become knowledge operators. Knowledge discovery in databases, traditionally also called data mining, introduces primary important step in technology of knowledge control.

Suitable source for discovering of nontrivial information can be data warehousing. In general some databases contain beforehand properly prepared data. Other technologies, look like OLAP, relate to analyses over multidimensional databases built up as the logical or physical layer over data warehouse (Schreiber, 2008). An application of data mining over multidimensional databases – OLAM (On–Line Analytical Mining) presents tendency of development which allows a qualitative new access for knowledge discoveries and their application by the determination (Halenar, 2011).

2. AREAS OF APPLICATION OF DATA MINING

One of the important steps is to identify the problems of production systems which can be solved by using the process of knowledge discovering in databases. These problems can be from the area of management production process, predict the behavior of the process, process optimization etc. The process of knowledge discoveries from management of the production systems can be applied to solve many problems. Here belong:

- identification of production parameters influence on a production process,
- identification of breakdowns in production process,
- identification of times series of final production,
- deviations (divergences) detection from plan during the production,
- failure states detection of production equipments,
- production quality assurance,
- production process optimization,
- workstations layout optimization,
- optimization of storage subsystem,
- prediction of consumption of production material,
- prediction of customer behavior,
- failures prediction in production process etc.

There are many problems that occur in the production process. It is important to choose the problem correctly and appropriate way of solving. The process of knowledge discovery in databases gives a lot of tools for solution of these problems.

3. PROJECT AIM AND STEPS OF PROCESS

The whole process of obtaining, storing and preprocessing data was divided into several steps. These steps are shown in Fig. 1. The authors expected that suggested procedure will discover new knowledge. Then the new control strategies will be defined on the base of the new knowledge.
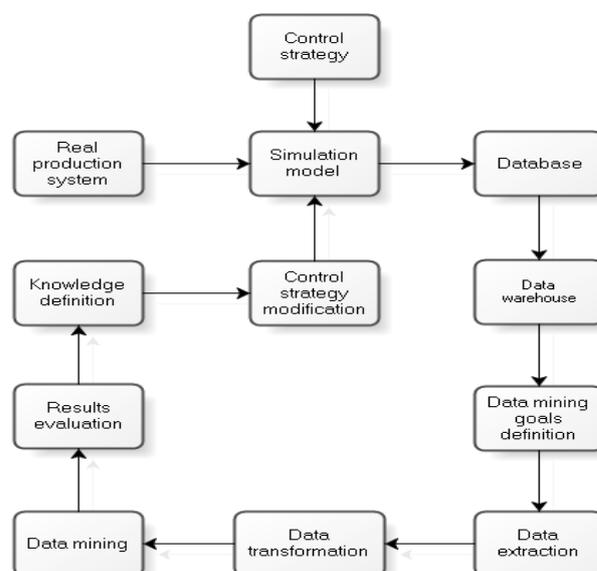

Fig. 1. The conceptual model of the project

4. SIMULATION MODEL

The simulation model in Fig. 2 consists of four workstations, where two parts are produced. The important production data are calculated during the simulation run. These data are stored into relationship database base on the Oracle 11g (Juhas et al., 2009). The authors have designed the structure of database and the connection between the simulation model and the database.
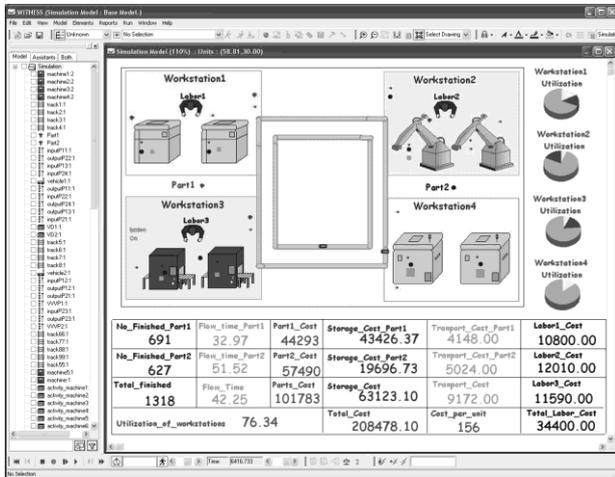

Fig. 2. Proposal of simulation model

5. DATA WAREHOUSE

The authors have decided to create the data warehouse for data preprocessing. The data warehouse has been designed and built by using Oracle Data warehouse Builder 11g.

During the ETL (Extract, Transform and Load) process has performed several data adjustment and data cleaning (see Fig. 3). The authors have verified whether the set of transformed data contains the data which values are significantly different from the other data, also called outliers. It was necessary not only to determine whether the data set contains outliers, but also the cause of their existence. It was necessary to identify the origin of such values and decide whether to include these values in the data set or removed from it. Similarly we have identified missing and noisy data. Thus prepared data set was transformed to the designed data warehouse.
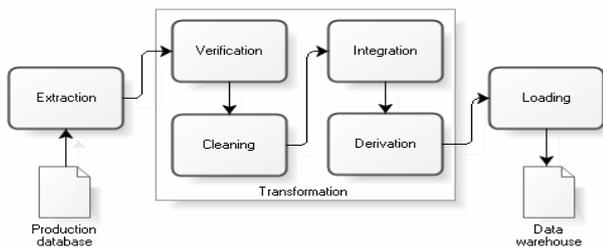

Fig. 3. ETL process

Multidimensional model of data warehouse in Fig. 4 is designed as a star scheme and uses the fact table and six dimensional tables.

Some data mining tools allow direct connection to the relational databases. Data obtained from relationship databases have to be cleaned and adjusted directly to the data mining tools. Therefore data analyses are complicated and lengthy. The data mining tool collects the preprocessed and fully transformed data from the proposed data warehouse. The whole process of data mining was facilitated and accelerated by creating the data warehouse.
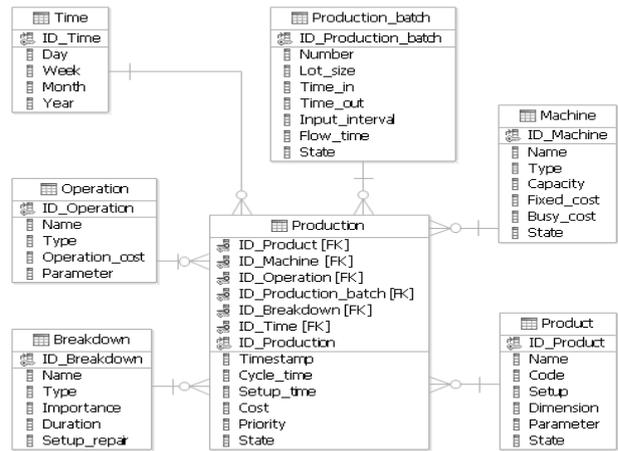

Fig. 4. Proposal of multidimensional model

6. THE FUTURE RESEARCH

The next logical step is implementation of knowledge discovery in the production systems. This solution will use to achieve a better understanding of a production system. The process will also use to gain new knowledge for predicting future behaviour of the production system. The new discovered knowledge will help managers in their final decision making.

The application of knowledge discovery process from databases in the production processes management will help to identify the impact of manufacturing parameters on the production process and the subsequent optimization of production process.

The future research will be oriented into the evaluation of the gained knowledge and its transformation into control strategies

7. CONCLUSION

The reached results cover the first stages of all knowledge discovery process and their implementation for new more efficiency of control strategies proposal of production systems.

8. ACKNOWLEDGEMENTS

9. REFERENCES

Fayyad, U. M.; Piatetski-Shapiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery: An Overview. In: *Advances in KDD*. Fayyad, U. et all. pp. 1-34, AAAI Press/The MIT Press, ISBN 02-62560-97-6, USA

Halenar, R. (2011). Contribution of Near Real Time ETL. *Proceedings of ICDDM 2011 Sanya, China.* ISBN 978-1-4244-9608-2, IEEE, Sanya

Schreiber, P. (2008). *Expert and database system cooperation in technical application,* Hochschule Anhalt, ISBN 978-3-86011-017-1- 1, Köthen

Juhas, M.; Juhasova, B.; Stremy, M. & Remenar, M. (2009). Non standard forms of data loss prevention in Oracle. *Proceedings of The 2nd International Conference on Advanced Computer Theory and Engineering. Cairo, Egypt.* ISBN 978-0-7918-0297-7, pp. 333-340, ASME, Cairo

Trnka, A. (2010). Market Basket Analysis with Data Mining methods. *Proceedings of ICNIT 2010 Manila, Philippines.* ISBN 978-1-4244-7579-7, pp. 446-450, IEEE, Manila