AN EFFECTIVE APPROACH USING BI TECHNIQUES TO ANALYZE PROCESS CONTROL PARAMETERS DURING STEEL PRODUCTS MANUFACTURING

Veena N. Jokhakar, Dr. S. V. Patel

ABSTRACT

Analysis is a process of understanding data more precisely. The process of analysis differs according to the data we are trying to analyze. Typical steel product manufacturing involves many processes running 24*7 in sequence comprising of different process units. These process units have sensors generating useful data for process control. These parameters have to be analyzed and monitored in real time to get a good quality product.

The conventional methods to analyze these data are inefficient. Hence, in this paper we present an approach for analysis that combines various process control and quality control parameters, uses online analytical processing (OLAP) system and provides flexibility to the user for necessary analysis.

Keywords: Data Warehouse, Online Analytical Processing, Statistical Process Control, Statistical Quality Control, Cube.

1. INTRODUCTION

Any company would always need to collect, organize and interpret the wide variety of information available in their business. Recent advances in computers and manufacturing techniques have made it easy to collect and store all kinds of data in manufacturing enterprises [4]. There is always a need to measure, understand and control the variables that affect the business process. Typically in a manufacturing industry, use of custom designed online analytical processing system (OLAP) for interpreting processing variables could be an effective way. As Quality is now playing an essential role to support manufacturers surviving in a competitive industrial market, to ensure that the final products fulfill the required standard with the minimum manufacturing cost and time [2]. For an example, the process of manufacturing large steel coils, involves large number of processing steps, each step requires strict Statistical process control and quality control and keep track of all details of the processes that went out-of-control or could not maintain required steel coil parameters.

In order to control such integrated process, we need to have a system, which is intelligent enough to do a great amount of data crunching and analysis on the data that the user selects dynamically. It needs to collect, organize and interpret the wide variety of information available during each sub process.

In the paper, we propose an effective approach using OLAP to achieve the goal. The proposed system will be using the various aspects of the SQL 2005 for the required purpose such as SSAS for building cubes, SSRS for presentation of analysis. It would create a dynamic and temporary cube on the fly for statistical analysis using charts like Histograms, Pareto Charts and different types of control charts.

One of the various components in data warehouse is information delivery component. And one of the various ways of information delivery is OLAP (Online Analytical Processing systems). There are various types of OLAP like ROLAP, MOLAP and HOLAP as discussed in [6]. Multi-dimensional OLAP (MOLAP) is where data is stored in a multidimensional cube. The storage is not in the relational database, but in proprietary formats. Relational OLAP (ROLAP) is a relational back-end wherein operations of the data are translated to relational queries. Hybrid OLAP (HOLAP) is the Integration of MOLAP and ROLAP. Desktop OLAP (DOLAP) is a simplified version of MOLAP or ROLAP. OLAP analysis on a multidimensional model involves selecting various slices of the data cube (i.e. fixing the value of a dimension by selecting a record from the corresponding dimension table) and aggregation of facts with respect to a dimension [1]. Some data mining tools are bundled with online analytical processing (OLAP) packages, while others are designed to run against relational databases or data warehouses [3]. Still others are extensions of standalone analysis packages and run against local data tables such as that found in a spreadsheet environment [3].

One another way of information delivery component is statistical analysis. This paper uses statistical process control and statistical quality control (SPC and SQC) charts. SPC is the primary analysis tool of quality improvement. It is the applied science that helps you collect, organize and interpret the wide variety of information available to your business. SPC can help you measure, understand and control the variables that affect business processes. SPC analyzes the variation in whatever process we are measuring: First, control charts demonstrate how consistently the processes are performing, and whether one should, or should not, attempt to adjust it next, the statistical process control chart compares the process performance to customer's requirements, providing a process capability index as an ongoing, accurate direction for quality improvement.

Finally, control charts and its resulting process capability index quickly evaluate the results of quality initiatives designed to improve process consistency.

SPC is one of the essential tools necessary to maintain an advantage in today's competitive marketplace.

The paper has been organized as follows: Section II describes the steel coil manufacturing process and highlights the importance of analysis of various data during each sub process. Section III describes different ways of data analysis Section IV shows proposed system for analysis with OLAP and SPC/SQC and Section V concludes the paper.

2. STEEL MANUFACTURING PROCESS

Steel manufacturing companies produces steel products like coils, plates, pipe etc. A typical manufacturing process starts from HBI/DRI (Hot-briquetted iron / Direct reduced iron), which makes bricks/balls of iron ore. This ore is then melted in electric furnaces named as EAF (Electric Arc Furnace). This melted iron is then converted into steel by adding many alloys. The liquid iron/steel is called as a heat. This liquid steel is then converted into steel slabs in caster and from slabs to coils in hot stripping mill (HSM). Further in Service center, there are many customized services and process that are performed according to the need of the customer like smoothening, galvanizing etc.. The whole process that was just discussed has been shown in figure 1. There are many units as shown in the figure 2. An example would be that a heat is made by melting iron ore, which gets converted into a slab weighing around 25 T of around 10 meters each in length, which gets converted into a coil of approximately same weight but 100 to 1800 meters in length. Now this coil can further be processed into small packs or be still elongated to almost 2500 meters to be converted into pipes for transporting gas, water, etc.

All these are produced with certain controlled parameters. But due to the dynamically changing parameters, sometimes these parameters are not adhered to. Also, each parameter may have a band, which may not be known to a user.

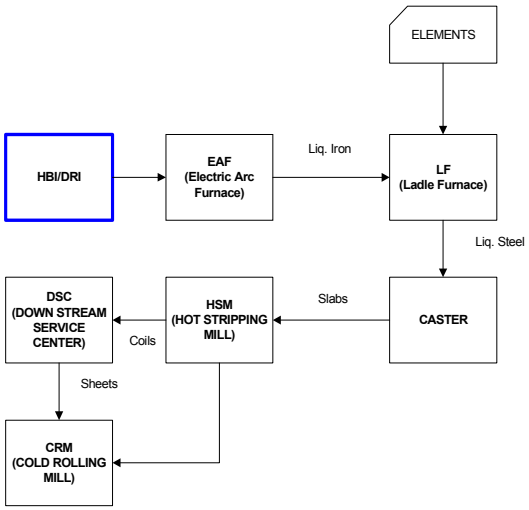

Figure 1: Steel coil-making process

The control parameters in each of the processes are being stored in continuous basis into different databases. And these data of controlled parameter are analyzed manually using SPC/SQC methodologies to find the processes that went out of control or were below the limits. Hence there is a need of having a different approach of analysis, which serves the purpose of following SPC /SQC and which could be user-friendly and extensible. Hence this paper makes an attempt to do it using effective approach.

3. SOME METHODS OF DATA ANALYSYS

There are different ways of data analysis, which can be followed. Some of them are as follows:

A.Using Excel: A simple method of statistical analysis is manual report generation in excel. However there are many problems with it. Human Error is also expected to happen leading to a big serious mistake in decision. People can mislead others by intentionally show wrong results to make process stable. One has to wait to get results as it is not automated system, nobody can see result any time. It is very difficult to do such a job manually as the data resides in different source databases.

B. Using OLTP: Online transaction processing system also allow analysis of data by report generation, but as is not user friendly as far as analyzing data is concerned as the user has to see the fixed format of report every time.

Moreover data as would be stored in RDBMS the data cannot be used for future enhancements for decision support systems and creating intelligent systems in future. Many such systems exist currently.

C.Using Data mining: Provides a way for data exploration & Analysis by Automatic or Semi-Automatic Means of Large Quantities of Data in order to Discover Meaningful Patterns & Rules. One of the implementations of DM techniques in manufacturing and their implementations on product design and manufacturing has been discussed in [4],[8].

D.Using OLAP: Online analytical processing systems is way of multidimensional analysis wherein the user is provided with various ways of data presentation like drill-down rollup up, multidimensional analysis etc. As mentioned earlier in this paper, there are different types of OLAP, like ROLAP, MOLAP, HOLAP and DOLAP. [2] Shows a data mining system (DMS), which is incorporated into emerging technologies including on-line analytical processing (OLAP). [5] Also shows a solution that urged the transform that was from Traditional Information Management System to the Information Analysis Decision System.

E.Using Statistical analysis: Statistical analysis is one many ways of analysis of huge amount of data where the nature of data is numeric. [7] Shows a method and a system for controlling a manufacturing process statistical process control data or other statistical indicators of performance from the production process and specification data are displayed in tables.

As we see that though above approaches for analyzing data are good enough according to the amount and type of data, but cannot be used as it is for the analysis of continuous real time data which needs lots of different complex analysis with user friendliness. Hence the approach stated in the paper tries to combine the strength of OLAP and Statistical analysis particularly using process control and quality control parameters for real time processes depending on the users need.

4. PROPOSED SYSTEM FOR ANALYSIS WITH OLAP AND SPC/SQC

This system would help the company to improve the quality of products and performance of the plant It uses the basic concepts of statistics and implements them by using SPC and SQC (statistical processes control and Statistical quality control) with multidimensional analysis.

The system needs to do a lot of number crunching and analysis on the data that the lines from the plant provide and then the system uses the data for further analysis on the various measures.

The data is available in a huge amount. Every 5 seconds 1 record with more than 100 parameters is stored in one or more tables from all the lines in the plant. This leads to approx of 20000 records per day per line.

The data that has been stored needs to be analyzed by various dimensions with various measures with any amount of granularity that would allow the user to visualize the whole situation in a readable and user friendly manner, therefore the system uses graphs to display the huge analyzed data in a readable format by the use of charts or graphs.

The various steps to be followed are as follows:

A. Data Grabbing and View Selection:

• As per the data storage architecture of the company, all data of a particular LINE, such as Galvanizing, Picking etc are stored in their respective database like ORACLE or SQL server and in respective servers. View is the entity, which works as a data source for this system. It will capture data from production table. There are basically three types of views for each line.
• The data are stored in views as per the data availability for the line. For eg, some Line 3 may not have data related to delay. The data stored in view are of 3 types.

Process Details – these contain data of every coil of every 5 sec for all the controllable parameters of the coil as per Location.
Delay Data – These contain the delay data related to the coil
Summary Data – Contains the summarized data as per time.
Meta data of this views will be stored in one table which will have fields like id, complex, data source name, line, userid, password, catalog. So whenever any

new view or line comes end user has to enter information about it into table. User has to select one view for further analysis in the system.

B. Filtration and Query Builder:

Each view will have different fields i.e. parameters. The entire parameter list will be listed and user can choose more than one attributes.

1) Filtration process: Filtration operators will be provided according to the type of the attribute selected for analysis.

2) Query Builder: will build query based on the type of SPC/SQC charts suitable for the attributes and views selected for the analysis, which shall be acknowledged to the user.

C. Cube creation as per users request:

Cube is an entity, which stores pre-calculated data. In this case cube will store all the data related to different charts such as histogram, X-bar, Pareto etc as shown in figure 2.

1) Create Data Source: If already cube name is exist in database then it will be dropped and new data source will be created with the same name.

2) Create Data Source View: In this part first dimension is being identified from attribute selected by user. For every related attribute one data table is being created and data type is being checked if it is numeric then user is asked for chart selection.

If chart type user has selected is xbar then subgroup data table is created. Minimum, Maximum, Internalwidth for each attribute will be calculated and according to those subgroups will be created. Each subgroup will be attribute of subgroup data table. Keys for data table will be defined then fact data table and relationships will be created.

If chart type is histogram then range data table is created. Minimum, Maximum, Internal width for each attribute will be calculated and according to those subgroups will be created.

Each range will be attribute of subgroup data table. Keys for data table will be defined then fact data table and relationships will be created. If chart type is Pareto or data type is varchar then only simple data table without range and sub

group will be created.

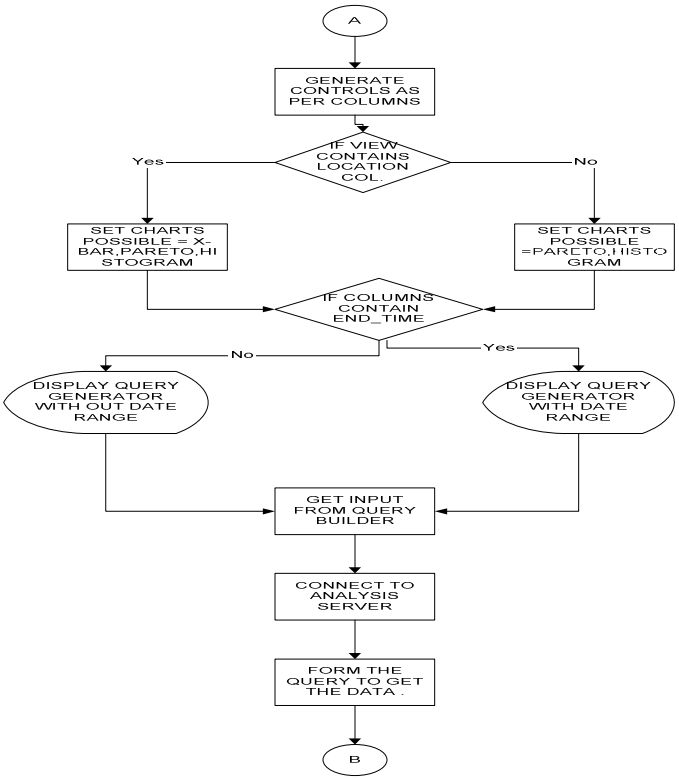

Figure 2: A chart showing user flexible approach

D. User Interface:

User has to select one dimension from the list. This list will contain all the dimensions in the cube depending on the chart. If Histogram is selected, then all the dimensions having ranges will be displayed.-In case of x-bar chart, all dimensions having the subgroup will be displayed -And in case of Pareto chart, all dimensions will be displayed. According to users selection.

1) Calculation for Histogram and Pareto chart:

Histogram charts has following calculation: Mean, median, min, max, skewness, sigma, UCL, LCL, Internal width.

2) Calculation for X-bar chart:

UCL, LCL, Sigma, min, max, range of all data, Average of Range, Range per data, Avg. per data. Various types of analysis can be done like Process is unstable, All

the data are correct,Some data are wrong, Some data are incorrect, Probably datas taken are over less period of time, Problem with some data.

By following this approach one may be able to perform multidimensional analysis on manufacturing company's process data and view the various control charts. This will also not occupy lots of memory as only the needed data is grabbed and analyzed. Moreover the cube structure may be deleted automatically or manually after few days of non-use.

E. Learning Module:

Learning module may also be added to help the user to analyze the selected data in detail and mine some interesting knowledge from the same. As we have the data in multidimensional manner and according to the statistics the user wants, mining on such a structure of data can give very interesting results. This may be the results of system selected algorithm or customized user selected algorithm according to the data.

5. CONCLUSION

The approach stated is a new concept and methodology for analysis that collects and analysis huge amount of data based on the users selection. Depending on the users selection the best suitable process control and quality control charts are selected and on the fly cubes are created for the data. This makes use of concepts of OLAP and Statistical process and quality control. The future enhancement for the proposal is the implementation of the same and that may be enhanced for data mining for predictions and providing remedies that form an Intelligent System/Or Agent.

REFERENCES:

1. Gergely Pintér,Henrique Madeira and Marco Vieira, István Majzik and András Pataricza ,Integration of OLAP and data mining for analysis of results from dependability evaluation experiments,2008.
2. K F Tsang, H C W Lau, S K Kwok Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, People's Republic of ChinaDevelopment of a Data Mining System for Continual Process Quality Improvement, 2006
3. Richard Kittler and Weidong Wang, Yield Dynamics, Inc. 2855 Kifer Road, Suite 100 Santa Clara, California 95051, DATA MINING FOR YIELD IMPROVEMENTS,2000
4. Kesheng Wang,Applying data mining to manufacturing: the nature and implications, Journal of Intelligent Manufacturing, 2007
5. WANG Cai,TANG Hui-jia, School of Information Science and Technology,Southwest Jiaotong University,Chengdu 610031,China), Application

research of manufacturing fault analysis base on warehouse and OLAP,2007
6. Yang Guang,Zhang Lei and Ai Bo, Beijing University of Posts and Telecommunications,Data Warehouse and Online Analytical Processing, 2000
7. Gary Fisher, Mineral Ridge, Ohio, Mark Clark, Mercer Force Company, Patent no. 54440478, 1995
8. Shakar kumar choudhari, Sunil kumar, Vinit K Mathur, Application of Datamining technology at Tata Steel, 2004

AUTHORS' PROFILE

Mrs Veena Jokarkar (M.Sc.(I.T.)) is a Assistant Professor at M.Sc.(I.T.) Programme, VNSGU, Surat. She has 2 years experience in industry and 2 years and 8 months experience in teaching. Her areas of interest are data ware housing, data mining, Information and security and Java. She has published 3 papers and she is actively involved in research.

Dr S. V. Patel (M E, Ph D) - Refer page number 29 for author profile.