1 7* International Mining Congress and Exhibition of Turkey- IMCET 2001, ©2001, ISBN 975-395-417-4

Knowledge Discovery in Mining Truck Condition and Performance Databases

I.K. Ataman
Department of Mining Engineering, University ofMissouri-Rolla, Rolla, MO 65409-0450, USA
T.S.Golosinski
Department of Mining Engineering, University of Missouri-Rolla, Rolla, MO 65409-0450. USA

ABSTRACT: Modem mining machinery is equipped with a large number of sensors that monitor its condition and performance. Data collected by these sensors is used to help with failure diagnostics, to warn the operators of impending failures, and to assess mine performance. Availability of large volumes of such data, gathered in real time and transmitted over wireless mine communication systems, together with availability of sophisticated data processing tools and the related hardware, provide a number of opportunities for further enhancement of mine performance. One such approach is application of data mining techniques for knowledge discovery in the databases containing condition and performance records of mining equipment. The paper describes the study undertaken to explore one such opportunity.

J INTRODUCTION

1.1 *Background*

Many newly developed technologies are being introduced to mining industry. Between these are the technologies that are capable of collecting variety of data on real time location, condition and performance of mining equipment. These technologies include Caterpillar's MineStar®, and in particular its part VIMS®, Modular Mining Intellimine®, and other. They were originally developed to facilitate development of a Total Mining System, a system that combines all mining related processes into one coherent and easy to control package, which allows optimizing the production and maximizing profit. Brief description of the selected technologies in question follows.

1.2 *Caterpillar's MineStar*

MineStar® is an integrated mining information svstem, developed by Caterpillar, Inc. and its alliance partners (Caterpillar, 2001). The system allows for tracking of machine health, productivity, machine and material movement, and drill management. It also includes Computer Aided Earth Moving System, CAES®, and an advanced truck assignment program. MineStar® has the capability of linking machines in the field to MineStar® office systems, as well as to other mine information systems. Caterpillar's alliance partners, Mincom and Trimble Navigation, have provided office software, and radio infrastructure software and GPS technology, respectively.

1.3 *Modular's Intellimine*

IntelliMine® is being developed by Modular Mining Systems, a subsidiary of Komatsu Mining Systems (Modular Mining, 2001). Applied in surface mines, IntelliMine® is to allow to maximize mine productivity by integrating optimized haulage fleet assignments, high precision GPS applications, Web reporting, and the latest spread spectrum radio communications network, IntelliCom®.

In underground mines IntelliMine® will maximize production and minimize equipment downtime by optimally allocating equipment, managing and monitoring the mining cycles, and improving equipment utilization

1.4 *Other Systems*

In addition to the two integrated, multi-module systems described above, a number of other, single purpose systems are available. These can collect, store and transmit a variety of data related to mining operations. Typical systems in this category include Komatsu's KOWA® (Komatsu Oil Wear Analysis), Cummins' CENSE®, Contronic Monitoring System® used by Volvo Construction Equipment Group and Euclid-Hitachi Heavy Equipment Ltd., and TRIPS® system used by Terex Mining.

1.5 *Mining Systems Innovations*

Introduction of innovative technologies to mining is a precondition to development of more efficient, computerized, total mining systems. At present the bulk of development is related to surface mines, the result of simpler communication set-ups and availability of GPS (Global Positioning Systems). For the time being underground mining lags behind these developments. It lacks cost-efficient and effective communication systems, and required reliability of data collection / transmission technology is lacking.

Total Mining Systems, when developed, are expected to comprise of (Golosinski and Ataman, 1999):

- Accurate and reliable data acquisition / transmission systems able to collect, transmit and process data in real time
- All computerized planning and scheduling systems mat interact with mine machinery to optimize the productivity and minimize human related delays and errors
- Vital signs monitoring systems that will track health and performance of every piece of equipment and facilitate timely predictive maintenance, which in turn will provide optimum availability and niinimize operating cost of each piece of mine machinery
- High-speed, broadband communication systems that can download megabytes of data and transfer it bi-directional without delay. This will allow the whole system to function in real time.

With introduction of systems like MineStar® and Intellimine® this dream İs about to be achieved. However, several problems still persist. One of them is lack of reliable predictive capability related to equipment condition and performance.

2 DATA COLLECTION AND STORAGE IN MINING

2.1 *Background*

The capability to both generate and collect data has been increasing rapidly in all industries. Today, countless databases are serving the needs of business management, government administration, scientific and engineering organizations, and many other applications. Resulting rapid growth in volume of data and database sizes has generated an urgent need for new technologies and tools that can intelligently and automatically analyze gathered data and transform it into useful information as well as extract knowledge contained therein. One such technology is data mining, discussed below. Data mining is becoming a research area of increasing importance.

Data mining, also referred to as knowledge discovery in databases, describes a process of non-tnvial extraction of implicit, previously unknown, and potentially useful information from data in a database. There is a number of algorithms that can efficiently mine wide variety of data. These include Decision Trees, Associations, Neural Networks, Clustering, Genetic Algorithms and Rule Induction (Berson and Smith, 1997). Based on these algorithms a number of software packages were developed, which by now are user friendly and allow to efficiently mine the databases. Some of the packages use just one algorithm and found applications in a specific areas, whereas others use integrated data mining architectures with multi-strategy approach and can be used in a multitude of applications. The latter include Clementine®, Darwin®, and Intelligent Miner® (Freitas and Lavington, 1998) between others.

2.2 *Knowledge discovery in mining*

With extensive number of various databases the mining industry is in a unique position to take advantage of the advances in information technology. In particular data related to equipment condition and performance, available in abundance and easily transmitted via the Internet, offers the huge potential for application of modern data processing tools. Research presented İn this paper is the world's first attempt to use data mining of these databases for knowledge discovery (Golosinski, 2001).

Subject of the investigations were databases containing mining truck performance and condition data collected İn De Beer's Jwaneng Mine in Botswana. The data was collected throughout the year 2000 from several CAT 789B off-highway trucks using Caterpillar's VIMS® system. IBM's Intelligent Miner® V6.1 software package was used to do data mining. The purpose of investigations was to confirm the possibility of finding useful information in the subject databases by applying representative statistical and other data mining techniques. Brief descriptif n of the tools used in the investigations follows.

2.3 *VIMS® (Vital Information Management System) of Caterpillar, Inc.*

Caterpillar's Vital Information Management System (VIMS) is a tool for machine management that provides operators, service personnel and managers with information on a wide range of vital machine functions (Caterpillar, 2001). It collects the data generated by numerous sensors that are integrated into the vehicle design, and processes this data into a set of equipment performance indicators. If VTMS detects an impending abnormal condition in any of the machine's subsystems, it alerts the operator and instructs him/her to take an appropriate action. This may be modifying machine operation, notifying the shop of needed maintenance or performing a safe shutdown of the machine. Use of VTMS improves availability, component life and production while reducing both repair cost and the risk of a catastrophic failures. A standard feature on large Caterpillar's mining trucks and wheel loaders, VTMS also provides production and performance information. Caterpillar's VIMSwireless® is a next generation technology that allows to transfer VTMS data through a wireless link between the truck and the office, and analyzes that data on an individual truck or a fleet-wide basis

VTMS System collects and records data in several files, namely: Event List, Snapshot, Data Logger Data, Trend Data, Cumulative Data, Histogram Data and Payload Data.

2.4 *Intelligent Miner of IBM*

IBM Intelligent Miner for Data (International Business Machines, 2001) helps to identify and extract high-value business intelligence from data, including high-volume transaction data generated by a pomt-of-sale, ATM, credit card, call center, or an e-commerce, to name just a few applications. This software package empowers its users to discover patterns, which might otherwise be unobserved, across volumes of data they would not be able to penetrate with other analytical tools. Intelligent Miner offers the opportunity to provide support to the mining process, as well as application services that enable development of customized software applications.

Intelligent Miner utilizes following statistical and data mining methods for knowledge discovery.

- Mining Functions: Associations, Clustering, Sequential Patterns, Similar Sequences, Classification, and Neural and Radial Basis Function Prediction.
- Statistical Functions: Bivariate Statistics, Linear Regression, Principle Component Analysis, Univariate Curve Fitting and Factor Analysis.

3 APPROACH

3.1 *Data Source*

As noted above, De Beer's Jwaneng Mine in Botswana provided all the VIMS data used in this research. The data was available as a number of sets, each set containing a complete record of 30 minutes of single track operation. The mine does not use VIMS Wireless, thus the limit on the length of data records.

Each 30-minute set was downloaded from individual trucks into a notebook computer and subsequently converted into the database format (MS Access) using VIMS office software VTMSpc®

3.2 Data Cleaning and Preparation

To prepare data for data mining it needs to be reviewed and cleaned to minimize the "noise", which could have detrimental effect on evaluation of the results. Typical problems that need to be addressed at this stage include missing values, empty columns, mismatch of various records, and the like. The data sets that covered full truck cycle, usually around 25 minutes for the investigated mine, were selected for further evaluation. Sets with less then 25 minutes of data are automatically rejected.

Selected data sets were cleaned and care was taken to make sure that identical parameters were included İn identical columns and that all the cells contained numeric values. MS Excel was used as data cleaning platform. In addition the column headers were matched with related columns using the MS Access Database.

The cleaned sets of data were then transferred to IBM's Intelligent Miner (IM) for statistical analysis and data mining. Sets of data can either be transferred to IM as flat files (MS Excel) or as database files (MS Access). IM also allows users to preprocess the data before using any data mining functions.

4 INVESTIGATIONS

4.1 *Correlation Coefficients*

The first step of the investigations involved search for correlation between various parameters. Knowledge of the correlation is a prerequisite for building a predictive model able to describe future performance and condition of the piece of equipment in question. Representative results of this work are shown in Figure 1 below.

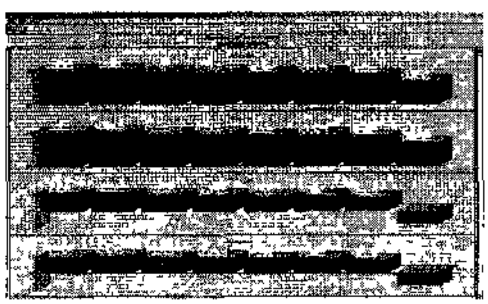

Figure 1 Sample Correlation Coefficients of Parameters.

Rather than seeking correlations between all parameters, the clusters of related parameters were defined based on the understanding of how the piece of equipment operates, and analyzed. The clusters of related parameters are more representative of equipment condition and performance that is a complete set of parameters, many of which are not related through common features of design or operation of the equipment.

The investigations concentrated on analysis of data related to performance and condition of truck engine, transmission and differential, which together constitute the most important components of the mechanical truck. Only those VIMS parameters that represent recording of sensor outputs were evaluated. The other, calculated VIMS parameters, derived by combining output of various sensors were excluded from further evaluations.

5 RESULTS

5.1 *Parameter Relations*

The data sets contained records of 85 truck parameters recorded at 1-second intervals. Several of the parameters were recorded in a binary form, and not all parameters were recorded in all data sets. From all these only the data, which was thought to be indicative of truck engine, transmission and differential condition was selected for data mining.

The longest data set that contained four 30-minute data sets for a specific truck, total of 2 hours of data, was then data mined. Unfortunately, not enough complete data sets were available for all trucks for which the VIMS data was provided. The incomplete sets were later used for confirmation and validation of the results of data mining.

Two parameters: engine speed and fuel flow were selected to describe and define engine condition (dependent variables). Both depend on and are related to a number of other parameters (independent variables), including several parameters calculated internally by VIMS. The latter were not included in the analysis for die reasons justified above.

Engine Speed and Fuel Flow parameters are the related to the highest number of other parameters (Fig. 2). The results also indicate existence of a relation between "Engine Oil Pressure" and "Engine Speed", as suggested by the VIMS developer. In addition, a strong relation was discovered between "Engine Coolant Temperature" and "After Cooler Temperature.

The other VIMS parameters do not show statistically significant relations with the two main parameters: Engine Speed and Fuel Flow Rate.

The results presented above were obtained by mining of data collected at one mine site only and using limited data sets (2 hours of data at most, with no more than 30 minutes of continuous data). It is, therefore, difficult to draw general conclusions applicable to the mining industry as a whole. However, the findings do apply to the operation in question.

	Correlation Coefficients	
Eng Speed	Related Para.	Cor. Coeff. x 100
	Eng Oil Pres	66 82
	Trn Out Spd	81 69
	Ground Spd	81 79
	Gear Code	82 55
	Gear Seiet	78 74
	Actual Gear	75.2
	Diff Lube Pres	67.43
Fuel Flow	Related Para	
	Eng Load	96.05
	Trb Out Pres	95.3
	Boost Pres	95.26
	ThrtPres	89.65
	Rl Exh Temp	78.61
	Lt Exh Temp	85.32

Figure 2 Correlation coefficients between Engine Speed and Fuel Flow versus related parameters.

Larger data sets are being acquired at present. It is hoped that their analysis may confirm relevance of the relations defined in this study to the industry as a whole. As this data is collected at several different mine sites, it will also allow to define site-specific relations between various VIMS parameters, if such exist.

5.2 *Effects of truck Payload on VIMS Parameters*

The next stage of the data mining was intended to analyze the relation between the payload carried by a truck and values of various parameters recorded by VIMS. For this purpose, a number of VIMS data sets were selected, each containing at least 25 minutes or continuous data. Out of 105 VIMS data sets that were available 35 fitted this criterion and were used in analysis. For all the selected data sets the data recorded for empty truck were combined with those for loaded one in order to define the relations between truck load and other VIMS parameters.

Since all data was recorded on the same model trucks, and since focus of investigations was on qualitative relations rather than quantitative ones, the external factors that may influence VIMS parameter values were not given consideration. These may include weather, road condition, type of hauled material, operator behavior and the like.

Overall the total length of the recorded data used in this analysis was 14,000 seconds (3.8 hours) with each of the 85 VTMS parameter values recorded once a second. The data was collected between January and October 2000.

The factor analysis module of Intelligent Miner was used to data mine these records. For compatibility with previous investigations the same parameters as before were selected as dependent and independent variables.

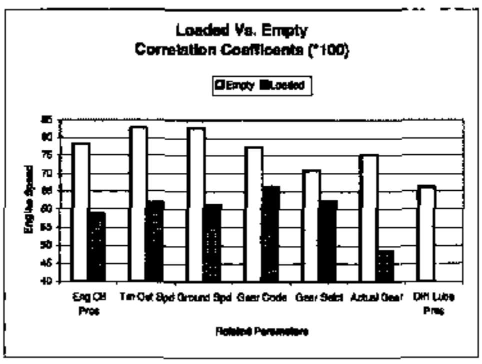

Figure 3. Loaded vs. Empty Truck - Engine Speed vs. Related Parameters.

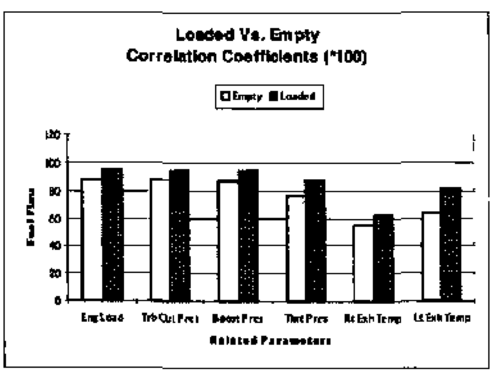

Figure 4. Loaded vs. Empty Track - Fuel Flow vs. Related Parameters.

The representative, selected results of data mining reported in this paper can be summarized as follows:

- If the truck İs empty, Engine Oil Pressure and other listed parameters are more strongly related to Engine Speed than when the truck is loaded. This relation is shown in Figure 3.
- When the truck is loaded Fuel Flow parameter is more strongly related to the listed parameters (in the referred figure) then when it is empty. This relation is shown in Figure 4.

Other results and relations are not discussed here for the sake of brevity.

5 CONCLUSIONS

Results of the investigations confirm that the data mining in general and IBM's Intelligent Miner in particular can be used to discover knowledge in the databases that contain data related to performance and condition of mining equipment.

The VTMS data is suitable for data mining with Intelligent Miner, providing that it is pre-processed to a format that is acceptable to Intelligent Miner. However, conventional VIMS System has a 30-minute on board data storage limit that prevents continuous data collection. Thus continuous analysis of data İs not possible with this system. VIMS wireless needs to be used in future work to carry the analysis on a continuous basis m Intelligent Miner platform.

Intelligent Miner was not originally developed to data mine mining related databases. As a result communication problems exist between data acquisition systems commonly used on mining equipment and the Intelligent Miner.

The character of data collected by various mining data acquisition systems varies. As a result the specific data mining algorithm that is best able to handle the data may vary. Matching the algorithms to the data types is the subject of the research that follows.

REFERENCES

Berson, A, Smith, S. J., 1997, Data Warehousing, Data Mining and OLAP, McGraw-Hill

Caterpillar Inc 2001, www cat com.

Chen, M, Han J., Yu, P S , December !996, Data Mining- An Overview From a Database Perspective, IEEE Transactions on Knowledge and Data Engineering, vol 8, no 6, pp 866-883.

Freıtas, A A, Lavıngton, S. H., 1998, Mining Very Large Databases with Parallel Processing, Khrwer Academic Publishers.

Golosmski, TS 2001 Dala mining for information Proceedings, 29th APCOM, Beijing.

Golosmski, T. S, 1999, Mining System Selection and Application Continuous vs Cyclical mining. Braunkohle - Surface Mining, vol 51, no 1,33-38, January

Golosmski T. S., Ataman I K., 1999, Total Open Pit Mining Systems: A Dream or a Reality?, Proceedings, 16th Mining Congress Of Turkey, pp. 89-92

IBM- International Business Machines 2001, www ibm com

Modular Mining Systems 2001, www.mmsi.com