Construction and performance analysis of data warehouse based on ship propulsion system

Qi Hu[1], Yutao Chen[1*] and Yuqian He[1]

[1]College of Power Engineering, Naval Univ. of Engineering, Wuhan Hubei, 430033, China

Abstract. A large number of operation data will be generated in the process of ship propulsion system operation. Aiming at the processing, analysis and application of these operation data, a data warehouse for propulsion system is constructed based on the operation data of a ship, and the multidimensional analysis and result display of data are realized by combining OLAP technology with visual analysis software tableau. On the basis of the data warehouse, the soft sensing of parameters is applied to the data warehouse. Based on the soft sensing data, the hydrodynamic characteristics of the propeller are studied, and the hydrodynamic characteristic curves of the propeller are obtained. The research of ship propulsion system based on data warehouse provides support for the comprehensive information of propulsion system data.

1 Introduction

With the continuous development of the shipbuilding industry and the continuous improvement of the ship information capability, the ship propulsion system has the ability to store equipment and system operation data for a long time in the operation process, forming a closed knowledge treasure house [1]. Due to the limitation of data processing and analysis ability, the operation data can not be used comprehensively, and the massive operation data forms "information island".

Aiming at the problem of insufficient processing, analysis and application of massive operation data, this paper constructs a data warehouse for soft measurement of propulsion system parameters based on the operation data of a ship's propulsion system, uses data warehouse ETL Technology for data processing, constructs the data warehouse model, and organizes the data under the subject of parameter soft measurement on the basis of data warehouse In this paper, OLAP on-line analysis and processing technology combined with visual analysis software tableau are used to analyze the hydrodynamic performance of propeller, and tableau is used to study the correction of propeller hydrodynamic curve by regularly updating data.

2 The construction of propulsion system data warehouse

2.1. Overall structure of data warehouse

The overall architecture of data warehouse can be divided into three layers, including source data acquisition, data preprocessing (ETL), data warehouse model design, online analytical processing (OLAP), and front-end tools [2-3].

As shown in Figure 1, The features of the three-layer structure are as follows:

(1). Bottom structure

The process from data acquisition to data preprocessing is the underlying structure of data warehouse system, which represents the process of data source and data ETL.

(2). Middle layer structure

The middle-level structure is the main body of the data warehouse. After the cleaned data is loaded into the data warehouse, the data warehouse carries out maintenance and management, determines the relevant topics, designs the model, and deploys the data after organization to form a data warehouse with corresponding topics.

(3). Top structure

The top-level structure is mainly to use OLAP analysis technology and front-end tools for data query and analysis, as well as to carry out some advanced applications of data warehouse. This part is directly oriented to users, providing users with the functions of data analysis, data visualization and data mining, so as to realize the comprehensive utilization of data.

The bottom structure and middle layer structure are

First author's e-mail:2013909115@ctgu.edu.cn
*Corresponding author's e-mail: yutao_jack_chen@163.com

the physical layer of data warehouse and the foundation of normal operation of data warehouse. The top-level structure is the operation layer of data warehouse, which is the basis of data warehouse function realization.
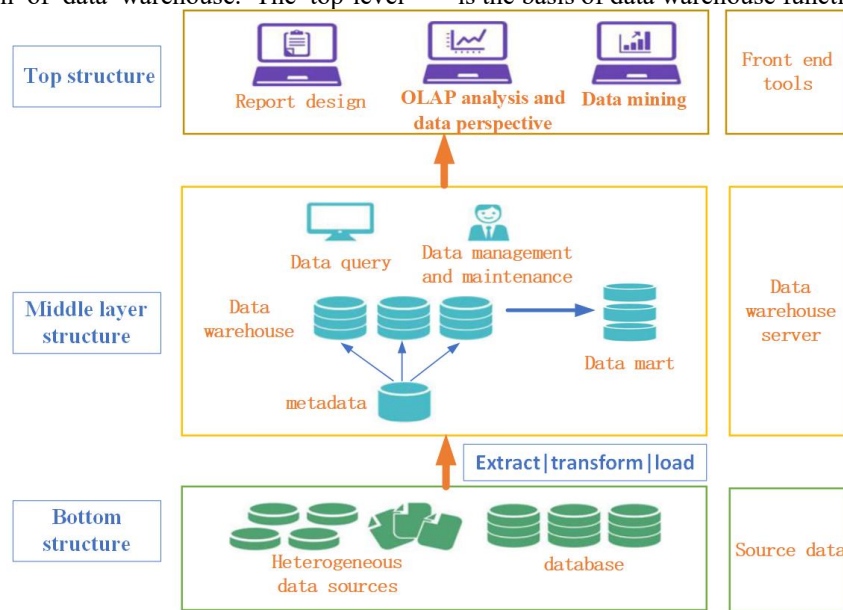

Figure.1 Data Warehouse System Structure

2.2. The realization of propulsion system data warehouse

The implementation of data warehouse is mainly aimed at the bottom structure and the middle structure. The implementation process mainly includes two parts: data preprocessing (ETL) and data warehouse model design. This paper takes Microsoft SQL Server as the design platform to implement the data warehouse of the propulsion system. The theme of the data warehouse is parameter soft measurement.

2.2.1.Data preprocessing

In the process of building data warehouse, the quality of data directly determines the quality of data warehouse and the results of top-level query analysis. Therefore, data preprocessing is necessary. Data preprocessing uses SSIS components in MSSQL to carry out ETL process. ETL is data extraction, data transformation and data loading. This process is the process of batch processing data and the key to building data warehouse.
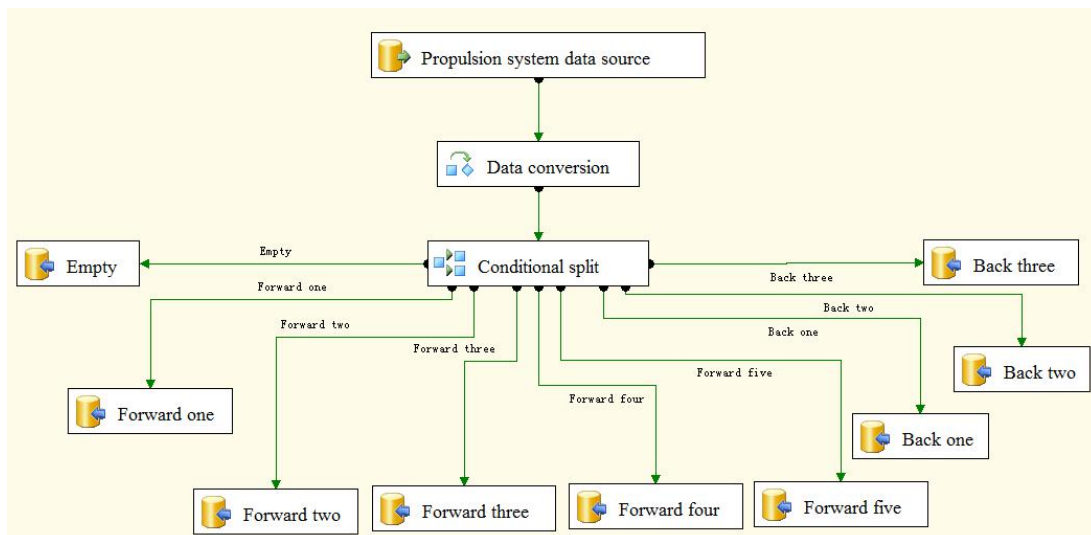

Figure 2. ETL flow chart of propulsion system fact data

As shown in Figure 2, the process of the actual data ETL of the propulsion system is shown in Figure 1. The data type and accuracy are converted through data conversion by connecting the data source of the propulsion system, and then the data filtering conditions are set according to the prior knowledge. The condition setting in the figure classifies the original data according to the speed and propeller angle under various working conditions, and the setting of conditions eliminates the noise data to a certain extent.

2.2.2.Data warehouse model design

The Data warehouse model design is divided into conceptual model design, logical model design and physical model design, which is the process of model transformation from relational data model and standardized model to multidimensional data model.

(1) Conceptual model

Conceptual model is the embodiment of the theme of data warehouse. The design of conceptual model often adopts the form of information package graph to expand the data cube, that is, the visualization of data cube is realized by using two-dimensional data table [4]. Information package graph refers to the relevant information data package established based on user requirements to meet the information requirements of data warehouse basic decision recommendations, including dimensions, categories and measures.

(2) logical model

The common logical models in data warehouse are star model and snowflake model. In the star model, each dimension table is directly connected with the fact table, and radiates from the fact table to the dimension table. One analysis topic corresponds to a star model. Star model is characterized by large granularity, strong data integration and high query efficiency, but it is easy to cause data redundancy due to no hierarchy [5]. The snowflake model is an extension of the star model. The model makes the data more detailed and standardized, reduces the data redundancy, and reduces the amount of data. However, it leads to the improvement of the model complexity and memory utilization, and reduces the efficiency of query analysis [6]. According to the characteristics of large amount of data and few

dimensions of propulsion system, star model is used as logical model.

(3) Physical model

The design of physical model includes the determination of storage structure, index structure, storage location and storage allocation, which affect the efficiency of query analysis after data loading. In fact, the construction of physical model is the realization of logical model on the development platform of data warehouse. The SSAS component in MSSQL is used to design the physical model.

3 Analysis and application of OLAP

3.1. OLAP analysis based on data warehouse of propulsion system

After the data loading and model construction of the system data warehouse are completed, OLAP technology can be used to analyze the data in multi-dimensional way. For simple query analysis, cube can be constructed directly through browser in MSSQL, and cube is displayed in the form of two-dimensional data table. As for propulsion system, the data browse box shows the operation data measurement values under the time dimension, ship dimension, working condition dimension and host dimension. The dimension attribute is set in the filter column to obtain the data to be queried. Meanwhile "+" and "-" respectively represent the drilling and rolling operations of data, and the row and column dimensions can be exchanged by dragging to realize rotation operation.
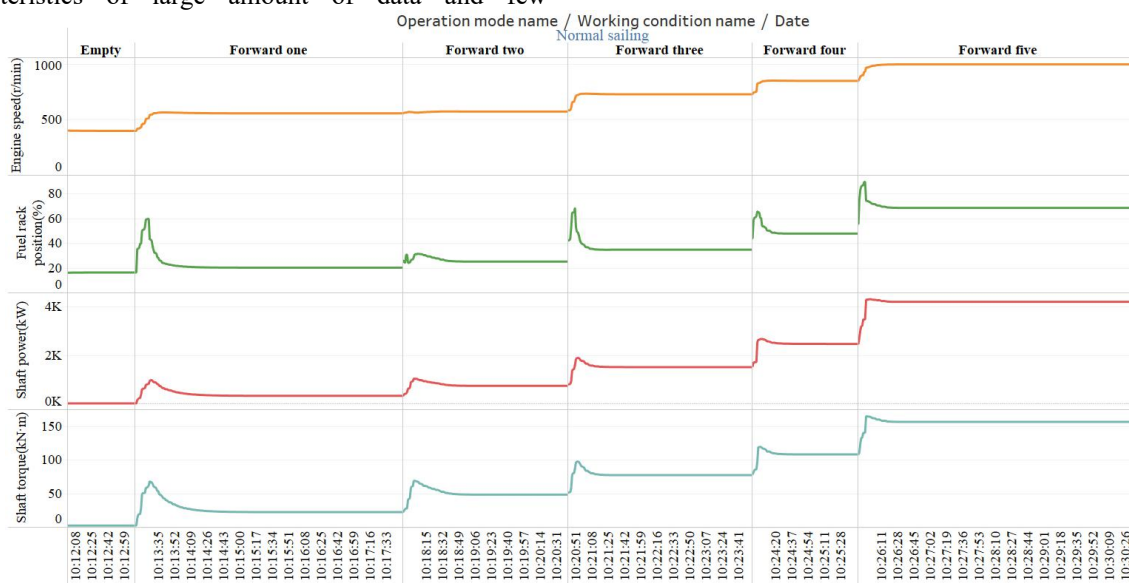

Figure 3. Multidimensional data analysis of propulsion system

For further analysis of the data, this paper uses tableau instead of Excel as a pivot tool. Figure 3 shows the result of multidimensional data perspective of a ship, which specifically describes the voyage from empty car to five working conditions, in which perspective dimensions are selected as working condition, operation mode and time dimension. The measurement values

include the right engine speed, fuel rack position, shaft power and shaft torque. It can be seen from the figure that when the operation data fluctuation basically occurs, the fluctuation time is within 120s, which indicates that the ship has good acceleration performance. In addition, the variation trend of parameters is the same, but the fluctuation value of parameter point differs from the

reference value due to different parameters. Compared with the reference value, the fluctuation rate of fuel rack position is 10% - 20%, while the fluctuation rate of shaft torque is 5% - 10%.

3.2. Application of data warehouse based on propulsion system

There are many operating parameters of propulsion system, but some of them can not be directly monitored. For example, there are more than 1000 data points of real ship operation studied in this paper, but some important parameters are not monitored by sensors. The soft measurement of operational parameters based on data warehouse not only provides more parameter points for propulsion system monitoring, but also reduces the monitoring cost of real ship. Taking the hydrodynamic parameters of the ship's controllable pitch propeller as an example, soft sensing is carried out, and new fields are created in tableau according to the formula.

$$K_T^{'} = \frac{T}{\rho \bullet D^2 \bullet (v_p^2 + n_p^2 \bullet D^2)} \tag{1}$$

$$K_M^{'} = \frac{M_P}{\rho \bullet D^2 \bullet (v_p^2 + n_p^2 \bullet D^5)} \tag{2}$$

$$\mu = \frac{v_P}{\sqrt{v_p^2 + n_p^2 D^2}} \tag{3}$$

Where is the thrust coefficient of the controllable pitch propeller, is the torque coefficient, and μ is the advance coefficient. Axial thrust unit n; shaft torque unit; sea water density, kg/m^3; shaft speed, r/s; propeller diameter.

In tableau, based on the soft sensing model of hydrodynamic parameters, the open water performance analysis of propeller can be carried out directly. According to the soft sensor model of the controllable pitch propeller, the hydrodynamic coefficients, μ are calculated when the pitch ratio is ± 100%, ± 93%, ± 50% and 0%. According to the relationship between and μ, the scatter diagram of the thrust coefficient and the advance coefficient μ of the controllable pitch propeller is obtained by tableau, and the curve fitting is carried out according to the scatter diagram, as shown in Figure 4. The thick line in the figure is the characteristic curve obtained by fitting, with both sides of the curve. They are 95% confidence interval of upper and lower respectively, and the fluctuation of data is basically concentrated in the confidence interval. The reading curve of historical data can be updated regularly to achieve the purpose of correcting the hydrodynamic curve of the real ship. The comparison of the new and old curves also provides a basis for propeller performance monitoring. Through the characteristic curve of the thrust coefficient of the controllable pitch propeller, the working condition can be reasonably planned and the pitch can be adjusted to realize the optimization of the operation state. The different pitch ratio provides a wider optimization space for the optimization, and the optimal operating point can be achieved through the combination of pitch and speed in navigation condition. If the speed is constant, pitch and speed are negatively correlated; if the speed is constant, pitch is positively correlated with speed and power.
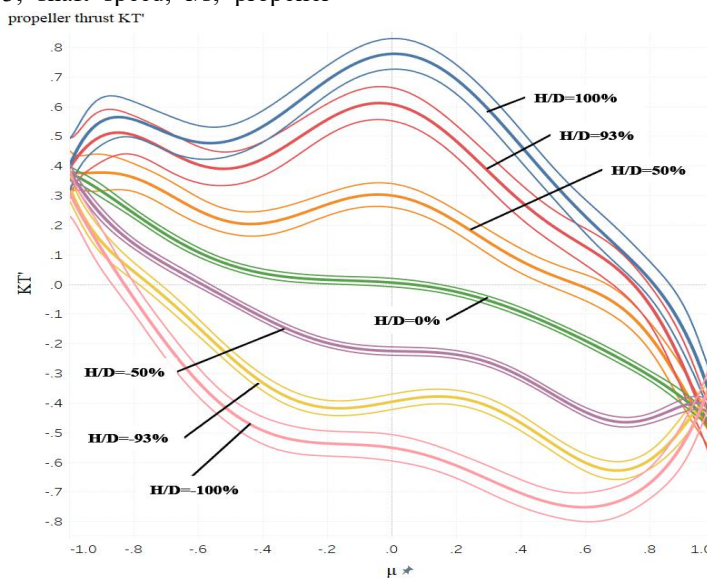

Figure 4. Thrust characteristic curve cluster

4   Conclusion

The application of data warehouse technology in ship propulsion system provides a strong technical support for the comprehensive information processing of ship operation data. The realization of data warehouse of propulsion system establishes the foundation for the application of ship propulsion system operation data management, system status monitoring and analysis, and also provides research direction for ship integrated information.

References

1. Liguo Yuan, hengwu Cui, Shijie Cao.(2016) Research on knowledge mining method of integrated management of ship propulsion system [J]. Ship science and technology,38 (23): 98-103.

2. Bimonte S , Antonelli L , Rizzi S . (2020) Requirements-driven data warehouse design based on enhanced pivot tables[J]. Requirements Engineering,(4):1-23.

3. Bonifati A , Cattaneo F , Ceri S , et al. (2001)Designing Data Marts for Data Warehouses[J]. ACM Transactions on Software Engineering and Methodology, 10(4):452-483.

4. Zaixun Guo. (2014) Research on the application of data warehouse and data mining technology in college enrollment decision-making [D]. Beijing University of technology

5. Elamin E, Alshomrani S, Feki J (2017)SSReq: a method for designing star schemas from decisional requirements. In: Proceedings international conference on communication, control, computing and electronics engineering, pp 1–7.

6. Qi Hu, Guoshi Zhu, Yutao Chen. (2020)Design of operation data extraction and analysis for ship monitoring system [J]. Shipbuilding Engineering, 49 (03): 112-116 + 120.