

Applying Knowledge Discovery System to Intelligent Manufacturing based on the Big Data

Qing Chen^{1, a}, Jun Pang^{2, b}, Yefeng Chen^{1, a}, Ming Yang^{1, a}, Yang Luo^{1, a}

¹Zhejiang Provincial Meteorological Information and Network Center, Hangzhou 31007, China

²College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430064, China

^azjhjxlchenqing@163.com, ^bpangjun@wust.edu.cn

Keywords: Knowledge Discovery System, Intelligent Manufacturing, Big Data.

Abstract. Intelligent manufacturing has increasingly become the major trend of the future development of manufacturing and core content. It is an important performance of a national comprehensive national strength. It is an important force to promote scientific and technological innovation, economic growth and social stability, and become the opportunities of the various countries' development and transformation as well as form a new competitive battlefield. And it is also the inevitable choice to build a new international competitive advantage under the new normal. However, in the development stage of the new things, rapidly, accurately and effectively in the first place to obtain new knowledge in the field of intelligent manufacturing new dynamic, is the best way to leading enterprises inevitable step in the market. So it is very important for the research of knowledge discovery system in the intelligent manufacturing under the Big Data. Based on the background of Big Data, this paper combines the knowledge discovery system and intelligent manufacturing field, designs the knowledge discovery system.

1. Big Data

In recent years, Big Data has attracted considerable attention. As early as 1980s, the American social thinker Alvin. Toffler praised the "Big Data" as "the third wave of the colour movement" in the book *The Third Wave*.

Big Data is a combination of several old and new technologies. It can help the companies to obtain meaningful data content, it helps to make deeper analysis of users' needs, and even potential knowledge needs of users, revealing the relationship between information resources, to provide more accurate knowledge discovery services. So Big Data can be thought as the ability of management of vast amount irrelevant data at the right time and applying it to real-time analysis and response. Big Data has three typical features: Volume (large amount of data), Variety (variety), Velocity (fast and efficient) ^[1].

In the era of Big Data, we can analyse more data, even sometimes deal with all the data associated with a particular phenomenon, rather than rely on random sampling. A well-known scholar at Harvard University suggested that the massive data sources had begun to quantify processes in all areas, whether academics, business or government ^[2]. With the rapid expansion of the amount of information, it has become the top priority of information services how to effectively use a large number of structured, semi-structured and unstructured complex data gradually

2. Intelligent Manufacturing and Big Data

The data have a huge value of innovation. We should not only pay attention to the actual amount of data, but also to the large data processing methods. As the complexity of the data increasing, the ability of transforming data into intelligence will be increasingly demanded.

How to achieve intelligent manufacturing from big data? A large number of entrepreneurs which come from world-class manufacturing have a general consensus. They realize intelligent

manufacturing from digital transformation. Digital transformation not only means that simply digital of companies, but also the core driving force of intelligent manufacturing which needed to be used to integrate the industrial chain and value chain ^[3].

In the field of manufacturing, one type of data is derived from the human trajectory generated data, the other is the machine automatically generated data. These two types of data constitute today's large multi-structured data sources. In the field of industrial data, we have to pay more attention to the integration of machine data and industrial data with human behaviour data in addition to continuing to care about human data or people-related data ^[4].

Enterprises can keep abreast of the processes, problems, and solutions in the production process, and to find new ways to create additional value based on large data tools, data analysis and mining.

Manufacturing industry can achieve business model changes, transform and enhance customer experience, improve the internal operational processes, keep abreast of industry trends in order to market decision-making step ahead by using big data tools and thinking.

3. Knowledge Discovery System Architecture

The data itself is worthless. The data will not let our manufacturing industry more advanced. It must be converted into information. And the information will be valuable for industry ^[5]. With the coming of the era of Big Data, users urgently need a simple and fast ideal platform which access to all the knowledge. Big Data is the basis of intelligent manufacturing, its core is the custom platform. If we don't invest in Big Data and data analysis, intelligent manufacturing pursuit of excellence operation will fall short. It is the Big Data analysis technology instead of Big Data itself to promote intelligent manufacturing, namely, the knowledge discovery system.

Knowledge discovery is non-trivial process of identifying valid, novel, potentially useful, and eventually understandable patterns from the data set. The process of Knowledge discovery translates information into knowledge, it find the data gold from the data mine. It will contribute to knowledge innovation and economic development.

In the field of intelligent manufacturing, through cooperation with enterprises build knowledge discovery service platform. In the industrial technology information, talent discovery, knowledge recommendation, and other aspects of service enterprises in the benefit analysis, customer relations, and so enhance the competitiveness of enterprises in all directions. The purpose of knowledge discovery is to break the limitations of the previous books directory, the full use of citation index and part of the literature, and to provide users with comprehensive and efficient knowledge mining and data analysis capabilities of the knowledge discovery system, in order to achieve the discovery from the resource to knowledge change.

Information organization supported by the technology of data mining, it carries on the knowledge association and the data analysis processing to the books, and further discovers hidden behind a large amount of data information, and then establish a powerful new generation of academic resource discovery platform to help the information users to obtain the required knowledge or node quickly and quickly.

3.1 Data Mining

Data mining is a process that it extracting implicit, unknown and potential useful information from a large amount of data which incomplete, noisy, fuzzy and random ^[6]. It is also known as knowledge discovery in database. It creates model and finds out relationships to make a decision and prediction among data from huge amounts of data by various methods and analysis tools.

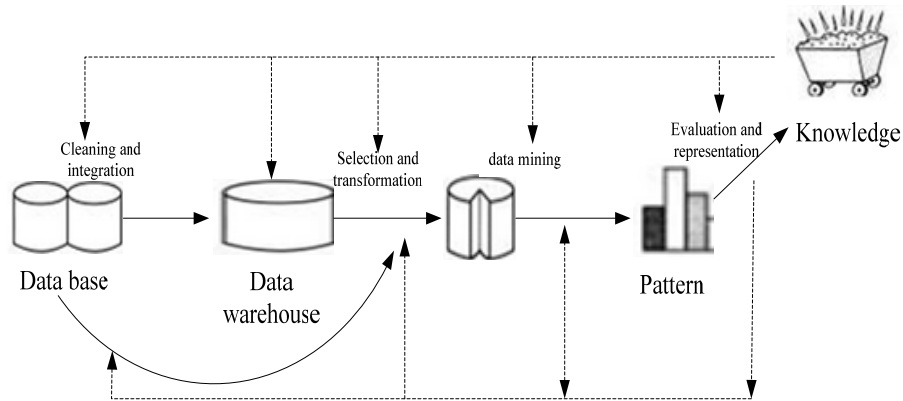


Figure 1. Data mining is a step in the process of knowledge discovery.

As shown in Figure 1, the data mining process includes the following steps: Understanding and definition of the problem: Data mining professionals work with domain experts to make an in-depth analysis of the problem in order to determine possible solutions and to evaluate the results of the study.

Related data collection and extraction. Collect relevant data according to the definition of the problem. In the process of data extraction, database query function can be used to speed up the data extraction.

Data exploration and cleaning. Understand the meaning of the field in the database and its relationship with other fields. Check and clean out the data which are contained in the data extracted from the data.

Data engineering: reprocesses the data. It mainly includes selecting related subsets of attributes to eliminate redundant attributes, and according to the knowledge discovery task to reduce the amount of learning, and transforming the way of expressing data to adapt the learning algorithm, in order to make the data and tasks to achieve the best match. This step may be repeated several times.

Algorithm Selection. According to the data and the problem which will be solved, select the appropriate data mining algorithm. And decide how to use the algorithm on these data.

Run the data mining algorithm. Using the selected data mining algorithm to extract the processed data pattern.

Evaluation of results. Evaluation of learning results depends on the issues which need to be resolved. The novelty and effectiveness of the patterns will be evaluated by domain experts. Data mining is a basic step in the knowledge discovery process. It includes a specific mining algorithm that is discovered patterns from the database. The knowledge discovery process uses data mining algorithms to extract or identify knowledge from a database based on specific metrics and thresholds. This process includes pre-processing of the database, sample partitioning and data transformation.

The task of data mining is finding models from data. It extracts useful and interesting knowledge and models from a large number of data by means of theory, methods and tools related to the development. The enormous data in a database often contains high-level information or knowledge such as rules, laws and assertion. We can't acquire this information only through the query process. Based on the data provided by the database, data mining looks for some inner relationship between data to find potential and important roles for forecasting and decision-making behavior patterns through data analysis and reasoning, finally it establishes a new business model to achieve the goal of help decision makers to make the right decisions.

3.2 Data Analysis

Based on the knowledge organization and presentation of existing structured and unstructured data, utilizing data mining and learning technology, reasoning knowledge discovery system could obtain change direction and trend of literature resources. Automatic and intelligent analysis can help users to get dynamic, informative, and advanced knowledge of literature. The primary role of data analysis is to help people to sample, extract, analyse operation process and check the analysis results by setting the human-computer interaction interface and utilizing software environment. The potential

regularity in huge amounts of information resources and its development trend will be shown dynamically and intuitively by means of visualization technology. Data analysis flow chart is shown in Figure 2.

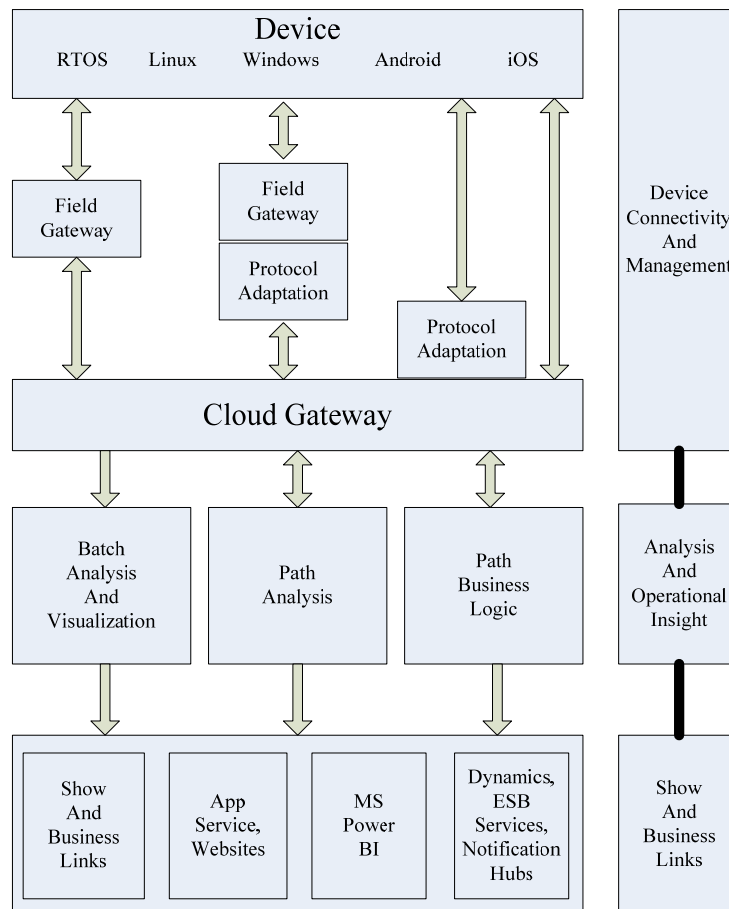


Figure 2. Data analysis flow chart

In the knowledge discovery system, data mining and data analysis make deeper development and application with information under cloud computing and Big Data environment. Data mining is the basis of data analysis; data analysis is the deepening of data mining. The content of data mining is the premise and guarantee for data analysis, data analysis systems reasoning and judge the development trend and direction of knowledge based on results of data mining, thus put forward ground-breaking and prospective prediction, verify the depth and breadth of knowledge mining, and feed back to knowledge mining system for its self-improvement. Therefore, data mining and data analysis, as two modules of knowledge discovery system, supplement and interact each other.

3.3 System Design Principles

The design of knowledge discovery system should follow the basic principles of system development.

1. Structural integrity. The knowledge discovery system gathers a large number of structured, semi-structured and unstructured complex data. It matches, analyses and mines the information agency and user needs and search behavior. Thus it establishes a complete system architecture. And it widely and accurately reveals the multiple, three-dimensional relationship of the various types of data and the knowledge network link ^[4].

2. Advanced technology. The knowledge discovery system innovates in the technical route. This system can not only dock with information service institution's original full-text information system, but also provide users with optimal results sorting and knowledge push through the advanced warehouse management technology, indexing technology, data mining technology and so on.

3. Platform security. Knowledge discovery system establishes a complete and effective system security, platform backup and fault handling mechanisms to ensure the stable operation of the platform. Job and synchronization system ensure data operations safe and stable. The design of

hardware redundancy fully protects the daily data's security backup. Real-time monitoring module realizes on the use of the database storage space and data access observation and monitoring, and it will alarm in case of emergency.

4. System development. The knowledge discovery system adopts extensible development language and database type, such as Distributed storage and distributed indexes, and so on. The performance of the system and its scalability have been greatly improved. The knowledge discovery system also provides OAI-DP services and standard interfaces. It allows seamless interfacing with other full-text access systems of information organizations. It can facilitate the docking of other platform on the system and the secondary use, so that the utilization of information being improved, the system's life cycle being continued. The old and new system of succession and development have been achieved.

5. Interface friendliness. The knowledge discovery system obeys to the basic principles of interaction design. It is user-centric interface designed. Its purpose is to plan and describe the mode that the users access to knowledge, then describe and communicate the knowledge information to the users efficiently. It provides users a friendly operating platform, visualization results display and export services, and a good user experience. For the user, the knowledge discovery system is a simple, easy to operate, and user-friendly platform.

3.4 Knowledge Discovery System Model

The knowledge discovery system is a process that it extracts implicit, unknown and potentially useful information from a large amount of data which is incomplete, noisy, fuzzy and random. Its purpose is to shield details for users, extract significant and succinct knowledge from original data and directly report to users, so that it can provide knowledge discovery service for business manager and information organization. Generally, resource integration, knowledge discovery and achievement exhibition are functional goals of knowledge discovery system.

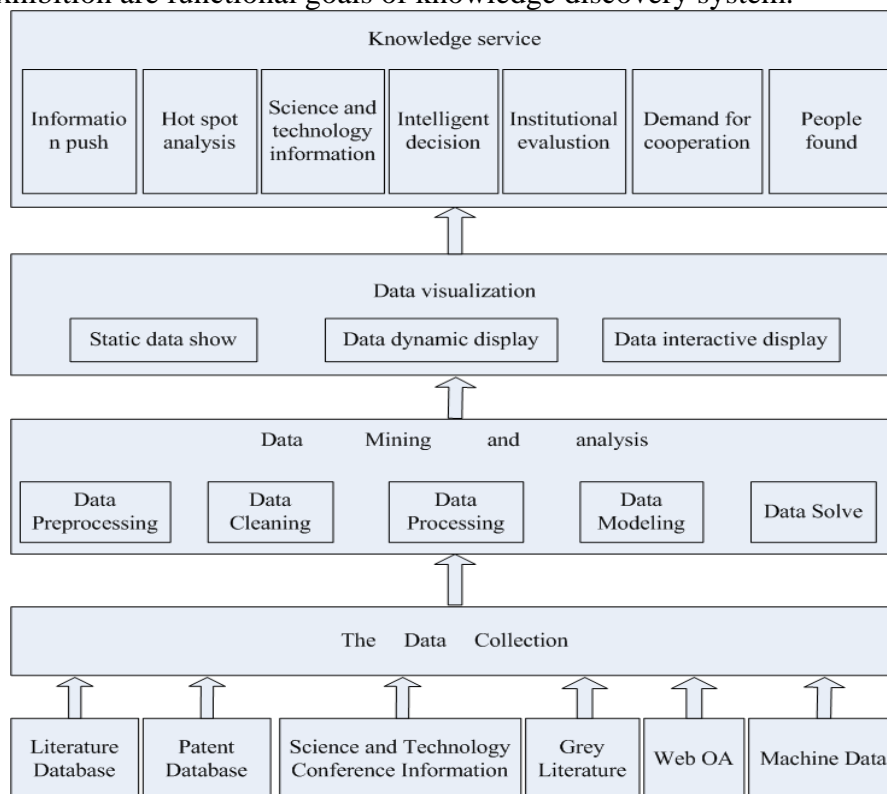


Figure 3. Knowledge Discovery System Framework Diagram of Intelligent Manufacturing Domain

The frame diagram of knowledge discovery system in the field of intelligent manufacturing is showed in figure 3, it consists of four parts: data collection, data mining and analysis, data visualization, knowledge service. Data collection module collects and receives original data, transports to data mining and analysis module, in which important and interesting data is extracted,

data is reduced to covert to appropriate format. Finally the generated knowledge model is evaluated; valuable knowledge is integrated into corporate intelligent system.

This system combines data and knowledge discovery system, discovers the unknown correlation between them. It also makes more possible to break information island Big Data used in operation, new data source such as social media, internet of things and so on, the analysis and solution capacity of Big Data, which are overall integrated together, knowledge discovery system can provide decision and prediction for management.

The firm boundary is becoming more and vaguer in manufacturing industry. Subversive innovation is the most unpredictable external factor. Interconnection entirely changed rules of the business, business managers need to realize cutting-edge technology and its correlations, make use of modern enterprise architecture to refine enterprise, and obtain more efficient, intelligent, high interest serving product through digital supply chain.

4. Summary

The matching of knowledge discovery system and enterprise management is the key to the enterprise performance. In the era of knowledge economy, the competition between enterprises is the competition of knowledge, information acquisition and their application ability. Knowledge discovery can discover useful knowledge from the mass data of the enterprise. Based on the mature information technology, this paper establishes an efficient search, discovery, accumulation, communication, sharing and reuse of knowledge platform. This paper also makes the enterprise to realize the true meaning of the knowledge sharing and reuse in a wider range, promotes enterprises to achieve effective knowledge management, so as to enhance the core competitiveness of enterprises. The system architecture proposed in this paper provides a useful way for the development of knowledge discovery system. However, next step, we need pay attention to the specific technical implementation details of the system architecture and its application effect.

References

- [1] Zhang Lili. Research from Big Data to Intelligent Manufacturing [J].China Industry Review, 2016, (7): 66-71.
- [2] Liu Jiangling. Research on Knowledge Discovery System Faced in Big Data [J]. Information Science, 2014, 03: 90-92+101.
- [3] Wang Ning. Analysis of the Digital Library Knowledge Discovery System under the Big Data Background [J]. Library Work and Study, 2016, (4): 58-61.EPL, 84 (2008).
- [4] Shouhong Wang HW. Towards Innovative Design Research in Information System [J]. The Journal of Computer Information Systems, 2010, 51(1): 11-18.
- [5] Wu Sizhu. Analysis of Hot Spots in Field of Data Mining and Knowledge Discovery [J]. Journal of Intelligence, 2010, 29(7): 18-24.
- [6] Ouyang Weimin, Zheng Cheng, Zhang Yan. International Survey of Knowledge Discovery and Data Mining Tools [J]. 2001, 28(3): 101-108.