



Double-layered big data analytics architecture for solar cells series welding machine

Feng-Que Pei, Dong-Bo Li, Yi-Fei Tong*

School of Mechanical Engineering, Nanjing University of Science & Technology, Nanjing, Jiangsu, China

ARTICLE INFO

Article history:

Received 31 May 2017

Received in revised form 24 September 2017

Accepted 31 January 2018

Available online xxx

Keywords:

Big data analytics

Solar cells series welding

MapReduce

Quality evaluation

Real-Time monitoring

Batch process

ABSTRACT

The rapid and extensive pervasion of big data has enhanced the revolution of the society. A great interest has arisen in the past five years for mining and sharing the massive data. However, implementation of the big data analysis is facing many challenges, such as the storage, transmission, and computing. How to make the decision more intelligent in latency time becomes a crucial requirement for many researches. In this paper, a double-layered architecture ATWDP of online and offline analytics for solar cells series welding machine industry is proposed and the distributed and parallel computing system can handle the above challenges. The ATWDP offers an approach to analyze the data gathered from various sensing devices stably and efficiently. Some key implementation technologies in the system are discussed, especially the Hadoop Apache framework. The evaluation of service quality based on Support Vector Machine-Dempster Shafer Theory (SVMs-DS) and Spark is an application scenario to illustrate the mechanism of the ATWDP. And a data set is used to verify the rationality of the ATWDP in the storage and processing. Test results show that the ATWDP platform has a good performance and is a suitable solution for dealing with the big data of solar cells series welding machine.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The United Nations Environmental Program (UNEP) declares that the Cleaner Production is the continuous application [1]. The most popular one is solar energy [2]. It has been about one hundred years since the discovery of photovoltaic effect in 1839 by French physicist. Till 1946, the silicon solar cells are modernized by the researchers Russell Ohl [3]. Then, in 1954, scientists from the United States Bell LABS G. Person invented monocrystal silicon photovoltaic cells, which increased the conversion efficiency of silicon solar cells up to 4% and made it widely applied in the world. And the widely using of the solar energy is an integrated preventive environmental strategy.

And the energy crisis of 1973 makes it possible for the photovoltaic power generation to be applied in the civilian market. Since then, due to increased government's supports, the photoelectric conversion efficiency has been improved and the cost of production has fallen dramatically. From 1974 to 1997, the power generation cost of photovoltaic battery dropped from \$50 per watt to \$5 in the United States, Japan and other developed countries. In 2007, the cost is about ¥4 per KWH and in 2010 it has been reduced

to ¥1. Given the advantages and the low cost, governments have vigorously supported and encouraged the development of photovoltaic industry. Solar photovoltaic power generation industry is developing rapidly worldwide shown in Tables 1 and 2 [4].

The researchers of EU make a forecast: In 2030, solar power will be accounted for over 10% of the total energy consumption; while in 2050, the proportion will be more than 20% [5]. Over all, solar photovoltaic power generation is expressed to play an important role in the future energy supply [6].

Lots of surveys show that, the solar cells welding industry with sophisticated operation and control has always been one of the leading enterprises because of the advancement in automation and information. At present, the solar cells series welding machine system has been widely used with Enterprise Resource Planning systems (ERP), Manufacturing Execution System (MES) and some other information systems to gather the data from series welding production line which can output about 2 million cells monthly. By analyzing these data, the evidence-driven decision can be made. The approaches to the extraction, storage and computing for data standardization and sharing become the urgent needs for this Industry.

As the solar cells series welding enterprises widely apply advanced information technology to manage the general operation, the solar cells series welding machine system acquires a related data stream of in the manufacturing process by the ERP and

* Corresponding author.

E-mail address: tyf51129@aliyun.com (Y.-F. Tong).

Table 1
The total installed capacity of global pv system (MW).

YEARS	2000	2005	2006	2007	2008	2009	2010	2011
CHN	19	68	80	100	145	373	893	3800
EU	211	2416	3377	5373	10671	16289	29617	–
USA	139	479	624	831	1169	1642	2520	4150
JPN	330	1422	1708	1919	2144	2627	3617	4733
Others	7556	1041	1225	1349	1771	2111	3130	–
Toltal	1455	5426	7014	9572	15900	23042	39777	67268

Table 2
China and the global solar cell production (MW).

YEARS	2000	2005	2006	2007	2008	2009	2010	2011
CHN	3	146	438	1088	3238	5851	7707	12894
GLOBAL	287	1793	2561	4000	9070	12951	27400	37200

MES. The typical data includes welding temperature, welding zone length, positioning data, and etc. The continuous manufacturing process, various of sensing devices, and timely and efficient data transmission make the solar cells welding data increase with an exponential speed with the characteristics of Volume, Velocity and Variety (Known for 3Vs) [7].

(1)Volume. The large quantity of working in process (WIP), the big scale of facilities and the high frequency of the signal acquisition, make the dataset huge. The CHR150-M10S series welding production line, as an example, can output about 1.92 million cells in a month and produce 80 thousand pieces of data (the dimension is 309) one day. And the data size is 2.8GB monthly in one facility while there involves thousands of this kind of acilities in operations.

(2)Variety. The dataset contains many types, such as Datetime, Char, Float, Boolean, etc. While for the variety of dimensions, there are pressure, temperature, time, length, angle, pixels, and so on. And for the value, it is a very large data range. For example, the magnitude of wind knife pressure is 10^6 Pa while the robot arm adsorption pressure is 10^3 Pa; the magnitude of welding time is 10^{-3} s while the delivery time is 10^1 s.

(3)Velocity. The velocity is mainly manifested in the short cycle of data acquisition, the speed of processing, the time of response, etc.

In term of big data or MapReduce, the researchers have established the distributed computing framework successfully. While earlier works on MapReduce focused on either pattern recognition or web extracting keywords other than industrial applications [8]. Such as; in 2013 Jindal et al. [9] established a Learning-Based-Java (LBJ) [10] based on parallel natural language processing(NLP) system; and Charm ++ was used [11] as a programming paradigm. And in 2015 Nesi P et al. [12] developed a distributed system for crawling web documents and extracting phrases which were provided by implementing the Hadoop. Galletti and Papadimitriou [13] investigated how the ICA Sverige AB and Masoutis S.A. applied the big data analysis (BDA) to perceive and drive to enterprises make evidence-based decisions.

Despite above, more industrial scholars are interesting to implement the massive data distributed system. Rashid M M et al. [14] proposed a parallel implementation of the regularly frequent sensor patterns (RFSP) mining algorithm, called RFSP on Hadoop (RFSP-H), which used a MapReduce-based framework to gain further efficiency. Big data implemented in cloud was presented for developing a HiTune platform, a lightweight and extensible application for Hadoop dataflow-based performance analysis [15]. Auschitzky et al. [16] proposed an analysis of the issues on how big data improve manufacturing by utilizing advanced mining

analytics to make evidence-based decisions. The methods of decision making were illustrated by the analysis of several cases in the study. Ramesh [17] put forward a structure with five layers: the source, platform, warehouse, mining and applications by using Teradata to address the storage and computing of the massive data.

Above all, in modern manufacturing environments, vast amounts of data differ from industries to industries in fabric and computing. The construction with special need for data analysis platforms has emerged as the basic mechanisms in different industries. While the majority are based on the Apache Hadoop Distribution Platform [18]. Compared with other industries, series welding big data have some distinctive features. The series welding data do not have video data and images data [19], and it is more similar to the telecommunications [20] or electricity data [21], whose data is made up of some structured type, like time, string and text. This reduces the difficulty of the big data processing for both the stream and batch processing. Therefore, according to the characteristics of the solar cells series welding machine industry, this study establishes a large data analysis platform ATWDP to make the analysis and application of the large data in the series welding. The rest of the paper is organized as follow: in Section 2, the architecture and function of the ATWDP platform is described in details; the Section 3 presents the related key technologies used in this platform; the Section 4 takes the quality evaluation as an example to analyses the operation mechanism of ATWDP; the Section 5 experiments the performance of ATWDP platform by several HiveSQL queries; And in the last section, we conclude the paper and consider the future work.

2. ATWDP architecture

Based on the analysis of big data technologies and the manufacturing industry, this paper proposes a manufacturing service oriented big data processing system called ATWDP. And the proposed ATWDP aims to execute a generic solar cells welding application in the distributed architecture. The Apache Hadoop framework [22] (open to access) has been selected for the realization of ATWDP and can be plugged in the Manufacturing Execution System (MES). The data-driven system offers quite a lot services which cover the remote supply, real-time monitoring of equipment and product status, and knowledge discovery. The characteristics can be described as below:

Fabric: Various kinds of resources are encapsulated into ATWDP which includes sensor information acquisition system, databases and software resources.

User-level services: this offers real-time monitoring, early warning and some remote support. These knowledge discoveries, by analysis and processing of the real-time data and historical information, can help obtain optimized decision-making supports.

Connection: this provides a convenient connection between MySQL and HDFS, HDFS and Hive, Hive and MySQL (By using Sqoop) [23] for service nodes and users.

Application: If the database is relatively structured, by changing the fields, the ATWDP can be applied in most area of industries.

Considering the large number of data types, the high real-time capability in the industrial site, and the high-speed transmission, this study discusses a double layered method which combines both online and batch offline methods by processing the real-time data in Real-Time Layer and dealing with the large scale hypo-real-time data in the Batch Layer. This architecture gives the exhibition of each layer and the overall architecture highlights their responsibilities. What's more, the terminology of this paper is established in this section.

As shown in Fig. 1, a hybrid service-driven solution is illustrated. In the architecture, there are three sections: the sensors node, the

Download English Version:

<https://daneshyari.com/en/article/6923865>

Download Persian Version:

<https://daneshyari.com/article/6923865>

[Daneshyari.com](https://daneshyari.com)