Shipping Big Data Applications Solutions Research

Submitted by: China
Shipping Big Data Applications Solutions Research

APEC Public-Private Dialogue on Advancing Trade Facilitation & Supply Chain Connectivity through APMEN with Inauguration of Asia-Pacific Model E-Port Network
Dr. Xu Kai (Kermel)
8-21-2015

Contents

- New Requirements
- How To Deal With
- Some Initial Results
- Not Only For Shipping
BACKGROUND

Today, international trade and logistics is undergoing profound changes. On the one hand, e-commerce is changing the traditional business model, a large number of transaction information has been generated through the network; on the other hand, with the IOT(Internet of things), sensors distributed in all aspects of logistics and transportation are also keep generating large amounts of data. Therefore, government departments and enterprises have access to more data than in the past, people urgently need new methods to maximize the value of these data.

Shanghai International Shipping Institute, the place where I work, provides consulting and information services for the government and enterprises. We know the value and significance of the data in the strategic decision of the government and enterprises, so we have done a lot of basic work in these areas.

Part 1

New Requirements
CSDB is very good.
But the data it provides are from the traditional way.

When we search for "Shanghai", the 62 kinds of statistical indicators associated with Shanghai will focus displayed on a page, such as: Cargo throughput of China's coastal ports, Oil natural gas and their products Traffic, etc.

From the traditional point of view, the CSDB has reached a advanced level.
### What is changing in Big Data Era?

<table>
<thead>
<tr>
<th></th>
<th>Acquisition</th>
<th>Integration</th>
<th>Storage</th>
<th>Processing</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Past</strong></td>
<td>Investigation/Manual</td>
<td>Paper/Digital info</td>
<td>Low speed/Localisation</td>
<td>Processing batch</td>
<td>Statistics/Causal analysis/Sampling analysis</td>
</tr>
<tr>
<td><strong>Now</strong></td>
<td>Internet of things/A</td>
<td>Enterprise Service bus(ESB)</td>
<td>Real time/Distributed/Cloud computing</td>
<td>Parallel Computing</td>
<td>Correlation analysis/Full sample analysis</td>
</tr>
<tr>
<td><strong>new problems</strong></td>
<td>Information Explosion</td>
<td>Standardization</td>
<td>Data Center</td>
<td>Super computing</td>
<td>Data Mining</td>
</tr>
</tbody>
</table>

**Big data research is going to solve these new problems.**

### How Big Data is generated?

- **The Internet of Things**
- **Computer Interaction**

**Big Data**
Some data collection techniques we have.

- Human Animal
- Cargo Container
- Ship
- Car Truck Machine

- Wearable Device
- Smart bracelet
- Logistics black box
- Terminal system for ship
- Car network suite

What shipping data do we have today?

- Geographic Data
  - Electronic chart
  - Sea floor map
  - Shoreline bath resources

- Monitoring Data
  - AIS
  -吹送 satellite positioning
  - Shipping lane
  - Identification
  - VDR

- Statistics
  - Transport volume
  - Transport capacity
  - Port throughput
  - Shipping route congestion
  - Deck operation efficiency
  - Freight rate

- Business Data
  - Ship visa
  - Ship registration data
  - Rescue request
  - BL data
  - Customers clearance data
  - Shipping order

- Textual Data
  - Shipping regulations
  - Maritime case
  - Shipping contact
  - Customers clearance data

- Media Data
  - Port surveillance video
  - Ship photos
"4V" Features of Big Data

Volume
- More than 20 million AIS data being produced every day throughout the world.

Variety
- Format heterogeneity, including video, images, other unstructured data.

Value
- 20 million AIS data can be compressed to 7 million.

Velocity
- Freight has strong timeliness, it changes every day.

Part 2

How To Deal With
Data analysis serves business decision

Data is the specific form of information that can be used to save, copy and spread.

So the “dead” data is useless, only to analyze the valuable part of the information can be called “big data”; otherwise, it’s just “a pile of data”.

And the purpose of data analysis is to help company make the right decision in the business.

---

Big Data Driven Model

Quote from the book "Big Data Driven Shipping Informatization"
Data and information

Shannon (C.E. Shannon), the founder of information theory proposed the concept of "information entropy" in 1948 to measure the number of "information" a data contains.

A large amount of information is has been lost from the first hand data to the data we see.

We can not see the whole picture without big data

Change an angle
Focus on detail and structure
The conclusion has changed.

226cm
192cm (AVG)
157cm
Distributed and Parallel Computing

NOW

Data utilization is developed with information technology

Stage of Shipping Informatization

Information System → Information Platform → Cloud Computing

Database → Big Data → Online Data

Evolution of Data Utilization

Quote from the book “Big Data Driven Shipping Informatization”
What Shipping Big Data Research?

Storage
A distributed database can store hundreds of billions of mobile target monitoring data.
There are 350 billion AIS data in the database.

Processing
Key algorithms for data cleaning and compression.

Analysis
Multiple analytical models and key algorithms.
Concurrent computation environment based on Hadoop technology.
Experimental environment of data analysis based on R and Python.

Visualization
Tracks and Scene playback technology.
Various graphics report technology.
Some Initial Results

Typical domestic route analysis results

By analyzing AIS data, we can know the real voyage condition of each liner routes and calculate actual voyage and fuel consumption of each route. So shipping company can account cost accurately and tell owner the punctuality rate of every route.
Global shipping route fitting results

We can calculate the number of ships and the total capacity of each route, which provides data support for market analysis and helps to predict the trend of market price.

Statistics of port production index

<table>
<thead>
<tr>
<th>2014-09</th>
<th>&quot;Shipping &amp; Port Big Data Laboratory&quot; Research Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Solid Cargo Ship</td>
</tr>
<tr>
<td>Statistics range</td>
<td>Berthing Time</td>
</tr>
<tr>
<td>Statistics</td>
<td>Unit</td>
</tr>
<tr>
<td>Busan Port</td>
<td>28203</td>
</tr>
<tr>
<td>Tianjin Port</td>
<td>54416.9</td>
</tr>
<tr>
<td>Singapore Port</td>
<td>37472.6</td>
</tr>
</tbody>
</table>

We have calculated the number of arrived ship and time of berthing to have an objective understanding of port loading/unloading efficiency and berth utilization rate, which provides data support for optimization of port management.
Regional emissions and energy consumption analysis

We can measure the concentration spatial distribution of pollutant emissions and carbon emissions precisely, and the results are being applied to the research of establishing gas emission control area in Yangtze River Delta and Pearl River Delta.

Ships Behavior Recognition Model

Previous studies are based on the identification of ship behavior.
Big Data technologies help us find all moored events

Within 15 days, on a global scale, all the ship mooring points.

According to location we distinguish between anchoring and berthing Events

Ship track data will be converted ship into a sailing log.
Not Only For Shipping

For Multimodal Transport

**Generic Target Monitoring System, GTMS**

Unified data format: *Multi-target Packet, MTP*
Our research is scalable

In the course of our research, we reserve the room for future expansion.

Future data analysis system

Not only for shipping industry, it will face to global trade and logistics.
Thank You!

Any comments, suggestions or feedbacks will be greatly appreciated!

Dr. Xu Kai
E-mail: kernyu@gmail.com
Phone: +86 13816950360

Aug, 2015