A Distributed Data Model for Port Administration and Onboard Information and Service Management

Chia-Hung SHIH*, Nobukazu WAKABAYASHI** and Saburo YAMAMURA***

Abstract

Along with the appearance of distributed concept, traditional centralized management framework has been transferred to distributed management framework, especially in electronic commerce field.

In the technical implementation level, because relational database technology is already very mature and easy to be implemented in distributed construction, so nowadays electronic business enterprises mostly use relational database management system (RDBMS) to store and manage their data. Although, in fleet management, because of ships cannot access to centric database through general internet construction when they are offshore, the data belong to ships themselves have the necessity of independent self-management. But RDBMS cannot completely satisfy this objective because it is not easy to be divided.

In this paper, we propose to use Extensible Markup Language (XML) to construct ship data management model, so then it could be applied for fleet management and future port administration. The data model can be composed of many kinds of information, such as hardware (ship name, ship type, net tonnage, etc), software (schedule, service, passenger, etc) and freight (freight kind, freight charge, deliver country, etc). Multi-level data model can provide the ability for fleet inter-management and communicating function to government related departments for port administration.

Keywords: XML, data modeling, distributed management, synchronization, XML Schema

1. Introduction

Maritime transportation industry is a high-complex industry; a fleet management company has to manage various ships with different characteristics divided by their duties (such as merchant ship, passenger ship, tanker ship, cargo ship, etc). And within the ship, it should be divided into various departments (such as engine, deck, information, etc) depend on the functions. According to these reasons, introduction of the distributed management mechanism is necessary. But the construction of ship industry is different from ordinary electronic business enterprise; the biggest difference can be analyzed as below. Though every activity under electronic commerce management mode is dynamic, but departments that each activity belongs to are still static. On the other hand in fleet management mode, ships can be treated as departments within the whole construction. But in this case, ships are not static but need to launch out in indefinite alternation by actual requirements. While a ship is sailing offshore, the communication ways with overland are limited into quite a small range, it can just use radio wave as usual communicate way; recently, because of the advancement of technology, communicate with satellite is also available, but this technology is not widely popularized yet and so the cost is very expensive. So we can just treat it as an emergency communication way. Under this kind of situation, if we use a traditional data construction like RDBMS to store and manage

---

* student member  Kobe University (〒658-0022 5-1-1 Fukae-minamimachi, Higashinada-ku, Kobe, Japan)  hiroshi@es.maritime.kobe-u.ac.jp
** member  Kobe University  waka@es.maritime.kobe-u.ac.jp
*** member  Kobe University  yam@maritime.kobe-u.ac.jp
information, because of the relationships between tables in relational database are complex so then it is difficult to divide it into a separate ship. And also it could cause the discontinuity and inconsistency of data.

Since the Extensible Markup Language (XML) is adopted as a Recommendation standard by World Wide Web Consortium (W3C) at 1998, it has become a very import emerging standard for E-commerce because its flexibility and universality. Many software designers are actively developing new systems to handle information in XML formats. Some researches in navigation field have used it as a data exchange format for applications \(^\text{(1)}\) \(^\text{(2)}\), but they have not considered the issue about the integration of data between ship itself and Shipping company.

In this paper, we propose to establish the ship data model by XML. By the characteristic of XML’s tree-structure, we can treat a ship as a child node under the management of the fleet management company, and each child node has its own independent XML sub-tree structure. So then during the sailing, it can manage its own data structure independently.

Another important characteristic of XML is that elements and their attributes can be defined by user. In our research, we combine the classified resources and Role-Based Access Control model \(^\text{(3)}\) to provide users appropriate resource access permissions. While a ship is arriving in the port, we provide an efficient data synchronization mechanism between the ship and the fleet management company.

The rest of this paper is organized as follows: First, Section 2 describes the characteristics of XML and presents how to establish the ship data model. Section 3 describes the data update mechanism and synchronization procedure for fleet management. Section 4 describes the convenience and possibility of using XML Schema to replace recent on-line applying form in port administration level. Finally, Section 5 gives our concluding remarks.

2. XML Data Model

In recent electronic business modes, because of the popularization of internet, the amount of users from various countries have been huge growth. In order to handle with the diversification of information and the necessity of simultaneous access from huge users, the international enterprises mostly use relational database to store their data. The advantages of relational database can be described below:

1. Reduce the duplication of data; data is independent and easy to establish applications to manage it.
2. RDBMS can provide the maintenance for the security and the consistency of data. It is also more efficient than simple file management.
3. RDBMS provides the index function for data access. So then when the capacity of data becomes very huge, the speed of data access could be much faster than directly accessing to files.

Although RDBMS has many advantages, it is not so suitable for fleet management. The main reason is that different from ordinary enterprises, the ships under the management of fleet management company (could be treated as the departments within a general enterprise) are not fixed but have the necessity of launching out and arriving in. While ships are sailing offshore, it is difficult to get the information from centric database through general internet construction. If we want to allot related information to a ship before it launches out, it would be also difficult to achieve because of the tables of relational database are related with each other and they are difficult to divide. In order to solve this problem, we propose to establish the ship XML data model for fleet management, the detail will be described in next sub-section:

2.1 XML features

The Extensible Markup Language (XML) is a simple, very flexible markup language. Originally designed to meet the challenges of large-scale electronic publishing, its primary purpose is to facilitate the sharing of structured data across different information systems, particularly via the Internet.

The characteristics of XML can be described below:

1. Tree architecture: XML can structure the data and display it with a tree-structure. Also the unstructured data that relational database is difficult to handle can be translated to the Semi-structured Data \(^\text{(4)}\) by definitions of DTD (Document Type Definition) or XML Schema. So then the unstructured data could be much easy to be handled by applications.
2. Normal text format: pure text store format allows
users to use various free text editor software to edit a XML document.

3. Unicode support: for the international era of nowadays, the support of multi-language is much more important than before. This feature is especially useful for an international enterprise like Shipping company.

4. Self-define tag: this feature provides us the extensibility. XML provides a markup structured data construction; it allows users to define specific professional field tags by themselves.

5. Data validation: precise data structure and grammar promise that all XML documents could be checked by XML Parser and make sure of that the names, sequences and layers of elements analyzed by different XML Parsers will receive the same results.

6. Independence of data structure: data is separated with display format. XML is a markup language for describing data; it does not define the display format within itself. The display format will be handled by XSL (eXtensible Style Language) \(^5\). Because of the content and display format are separated, it is much easier to read and maintain for both human beings and computers.

Because XML has the characteristics described above, it could be a more suitable data management model than RDBMS. So in our research, we propose to use XML to establish ship data model for fleet management.

2.2 Fleet management

Fleet management can be seen as a tree-structure with fleet management company on top and various ships below under management, so we can display its structure visually by XML. Our data model treats fleet management company as the root node, ships under management as child nodes; and each child node has its own affiliated data model. When ships are berthing in the port, the data model is a complete XML tree architecture. But when a ship launches out, the partial tree data belong to the ship should be divided from the original XML tree. And while this ship is sailing, it should upgrades to be a temporary complete XML tree and has its own root element (shown in Fig.1). After this ship comes back to the port, based on the policies we predefined, we can proceed a synchronization between fleet management company and the ship by exchanging corresponding child nodes data easily.

![Fig. 1 Separable XML tree architecture](image_url)

2.3 Ship data modeling

In our research, we focus on cargo ships and passenger ship to establish the ship data model. In order to implement this model to various objective (such as procedures of arriving in, cargo clearance, immigration declaration, etc), the data structure should include the elements described below:

1. Ship basic information: use to identify ship itself. Including name, description, flag, official number, length, draft, agent, operator, owner, net tonnage, etc.

2. Navigation data: the navigation record data of all kinds of navigation instruments within the ship, including radar, magnetic compass, GPS, speedometer, chronometer, etc.

3. Schedule data: the schedule information of the ship, including port of provenance, date of departure, port of call, date of arrival, time and date of entrance, berth, final port of destination, etc.

4. Service meta data: the meta data of applications applied within the ship, such as watch sign-in management service.

5. Cargo data: the information of cargo transported by the ship, including marks and Nos., number and kind of packages, description of goods, gross weight, place of shipment, place of destination, etc.

6. Passenger data: the information of passengers in the ship, including name, nationality, date and place of birth, port of embarkation, port of disembarkation, passport No, etc.

7. Crew data: the information of the crews in the ship, including name, nationality, date of birth,
sex, number of sea men's book, rank, etc. It is similar with the passenger information except the information of duty.

Because XML allow us to define the elements and their attributes within the document, the usability is raised hugely. In our research, in order to classify the data and raise the efficiency of synchronization, we define two attributes, property and update. The sample can be described below:

<Element property="public" update="true" >

In the sample above, the "public" value of property means that everybody is allowed to access to this resource. The "true" value of update means that this element has been updated before.

Because of the essence of XML construction, depending on the actual situation, the number of code line in the XML data model could reach several thousands, so then it is not so efficient to read the whole file. In order to solve this problem and consider about the security at the same time, we adopt Role-Based Access Control (RBAC) model \(^{(6)}\). The essence of RBAC is that permissions are assigned to roles rather than to individual users. Roles are created for various job functions, and users are assigned to roles based on their qualifications and responsibility. With this way, the task of specifying user authorization is divided into two logically independent parts: one which assigns users to roles and another which assigns access rights for objects to roles. Users can be easily reassigned from one role to another without modifying the underlying access structure. RBAC is thus more scalable than user-based security specifications and greatly reduces the cost and administration overhead associated with fine-grained security administration at the level of individual users, objects, or permissions. RBAC model can be simply expressed as the form \((s,o,a)\), where \(s\) is a subject (user), \(o\) is an object (resource) and \(a\) is an access right (permission).

By using the RBAC model and the property defined in our ship data model, we can achieve the objective of distributed management within the ship. For example, passengers are only allowed to access the general public information such as the schedule data within the ship data model. This can not only maintain the security but also raise the efficiency by reducing the capacity of data reading. By controlling the permissions of writing in data, we can also make the information belong to each department be in charge by the chief of the department. Such as the chief engineer can renew the navigation information of all machines in the engine room, but the chief officer can only read these information. When a ship arrives in the port, if we allow the agents of government departments accessing to the ship data model, we can also assign the corresponding information by their duties to them. Such as the custom agent can access the related cargo's information to apply the cargo clearance procedures, the immigration agent can access the related passengers' information to apply the immigration declaration procedures. The image figure can be illustrated as Fig.2.

![Fig. 2 Corresponding sub-tree according to duty](image)

### 3. Data integration

During the sailing of the ships, the data (such as sailing schedule, navigation instruments records) could have the necessity that should be renewed because of various unpredictable factors (such as weather, hydrographic condition). In our research, we treat the ship as an independent individual. While the ship is sailing offshore, it should manage its data by itself, and proceeds the data integration with fleet management company when it arrives in the port. The old ship data that fleet management company has should be read only during the ship is sailing to prevent from destroying the consistency of data. Because we adopt XML to establish data model and treat the ship as a node, so then the most simple synchronization mechanism is replacing the old node data of fleet management company with the renewed ship node data. But this mean is not efficient when the capacity of data is huge, so then we propose a
more efficient data synchronization mechanism in this paper.

The data integration mechanism we propose should be divided into two parts. The first part is the data update during the sailing of the ship; the second part is the synchronization mechanism when the ship arrives in the port. We will describe them respectively in the next two sub-sections.

3.1 The data update during the sailing

During the sailing of the ship, there is the possibility that the data should be changed because of various unpredictable situations (such as the change of weather or hydrographic condition, engine problem, artificial factors, etc). In order to record the path of data updating and achieve an efficient synchronization later, we define an update attribute for XML document elements. Within our XML data model, if any information of node $N_{ij}$ is updated (i denotes the i-th layer of the whole XML tree, j denotes the j-th node of the current layer), then the update attribute of $N_{ij}$ is set to be "true", and at the same time pass the value of "true" to the update attribute of its parent node $N_{i-1, parent}$ (the direct parent node of $N_{ij}$ within (i-1)-th layer of the whole XML tree). If $N_{i-1, parent}$ is not the root node of the ship data model, then repeat the processes above. In this paper, we call this mechanism as upward parent node chain update. The algorithm of it is described as Fig.3; the image figure is illustrated as Fig.4.

**Upward parent node chain update algorithm**

```java
If (Node $N_{ij}$.update == true) {
    While (Node $i-1, parent$ != Node root) {
        Node $i-1, parent$.update = true;
        \(i = i - 1;\)
    }
}
Node root.update = true;
```

Fig. 3 Upward parent node chain update algorithm

3.2 The synchronization when arriving in the port

When the ship arrives in the port, in order to maintain the consistency of data, the synchronization with fleet management company should be applied. The procedure of the synchronization mechanism we propose is described below:

Step1. Read in XML Schema to get the structure of entire XML document, mostly the relationship between nodes in different levels.

Step2. Compare the data that fleet management company has with the data of the ship just arriving in.

Step3. If the data of the ship has been updated, then start to search the child nodes of the first sub-layer of the ship root node.

Step4. If the j-th node of the current layer is not updated, then ignore all sub-tree below it and move to (j+1)-th node.

Step5. If the j-th node of current layer is updated, then start to search the its child nodes of its first sub-layer.

Step6. Repeat step.4 to step.5 until the whole XML-tree has been searched completed.

According to the above procedure, we illustrate the data synchronization flowchart as Fig.5.
Fig. 5 Data synchronization flowchart

By combining the two procedures of Section 3.1 and Section 3.2, we propose an efficient XML data synchronization mechanism. In fleet management level, the ships can manage their own data structures while they are sailing offshore; and when they arrive in the port, the synchronizations with the fleet management company would be applied to maintain the consistency of data.

4. Potential implementation for port administration

When a ship arrives in the port, it should apply many complex procedures with various government departments. The main procedures of arriving in include the cargo clearance of the Custom, the immigration declaration of the Immigration, the quarantine of the Ministry of health, Labour and Welfare, the clearance inward of each port, etc. And each procedure need to fill in various applying tables, such as the cargo clearance need to fill in general declaration, cargo declaration, ship’s stores declaration, passenger list, crew manifest, crew list, etc. In order to simplify these complex procedures, Japanese government are proceeding the optimization project of the import, export, and clearance inward procedures. The objective is to combine the ongoing multi-systems (such as NACSS, port EDI, etc) to complete the procedures such as clearance inward that has to apply to multi-government departments with one single window and one time data input/transmission for international ships. So then it would be much more convenient for users and also it can reduce the cost and the applying time. For now, basically these applying procedures can be proceeded to input various data with web browser as the input interface through internet. Comparing with old artificial operation era, we can say that it is already a very convenient way. But if we can use the characteristic of the authentication of XML document, there are still some potential to further simplify artificial inputting processes. The possibility of implementation of the XML Schema (an important XML document authentication mechanism) for port administration is described below.

XML Schema is a standard that was approved as a World Wide Web (W3C) Recommendation on 2 May 2001. It can be used to express a schema: a set of rules to which an XML document must conform in order to be considered validation according to that schema.

XML Schemas express shared vocabularies and allow machines to carry out rules made by people. They provide a mechanism for defining the structure, content and semantics of XML documents. The XML Schema definitions for XML documents include:
1. The vocabulary: element and attribute names.
2. The content model: relationships and structure.
3. The data types: data entity format (such as string, boolean, integer, etc).

If a XML document can be valid by a specific XML Schema document, then we can say that this XML document conform every restriction for our purpose, and it is a valid XML document.

We would like to take the Kobe port EDI system as an example here. When a ship arrives in the port, it could fill in corresponding tables through internet to apply various services (such as licensing of berthing facilities, watering application, licensing of wharf site, etc). Each column in the electronic form has its own format restriction to prevent for errors of storing data into database from user miss inputing.

As an implementation sample of XML Schema, we take the on-line electronic applying form of Kobe port
EDI system wharf site license for example (shown in Fig.6). According to the format restriction of each column, we can establish the corresponding XML Schema that defines the data format of every column in the form (shown in Fig.7); and then we can establish a corresponding XML document with this XML Schema document.

Because the format restriction of XML Schema is very regular, so when the user transmit the applying XML document to corresponding government department, if it can pass the check mechanism of XML Schema, then the government department can consider that the content of this document is correct. There is no need for human rechecking and the data of this document can be directly storing into the database.

This method can reduce the waste of internet resource in one side; and in other side it is possible to simplify the applying procedures. Cooperating with the ship XML data model we propose in this paper, we consider it as a potential implementation for port administration.

![Fig. 6 Kobe port on-line wharf site application form](image)

![Fig. 7 Corresponding XML Schema definition](image)

5. Conclusion

In this paper, we adopt Extensible Markup Language (XML) to construct ship data model, classify data model by XML tree-structure characteristic. Each ship can be treated as a child node for fleet management level, this XML data model can separate its child nodes, and the child nodes can be operated independently.

Also, we define two attributes (property and update) for elements in this data model to provide the function of resource classification and the index of data updating. Cooperating with appropriate user permission management, we can achieve the objective of access control and efficient data synchronization.

For the future work, we have two research objectives. The first objective is to develop a system that can automatically generate corresponding XML documents by the predefined XML Schemas according to various on-line government departments' port administration applying forms. Using the XML data model we structure for the ship efficiently to reduce the time of artificial
inputing and the waste of internet resource. So then we can further simplify the applying procedures. The second objective is to raise the speed of reading XML data. Because of the essence of XML construction, the capacity of the XML file could be very huge even reach several hundred MB. So how to transmit the desired information to specific users efficiently is also an important issue that should be discussed.

References
(2) Kouhei Hirono, Development of the agent system to assist a navigator onboard and collect operational logs for shore based process management, The journal of Japan institute of navigation, pp. 1-10, 2003.3.
(7) The optimization project of import/export and port/air port related procedure operation system, 16th Chief Information Officer meeting conclusion, 2005.12.28 (in Japanese).

Questions and Answers
Osamu AMAI (Electronic Navigation Research Institute): In your summary paper, you mentioned that the cost of satellite communication is very high. How much is the actual cost about it?
Chia-Hung SHIH: Thank you very much for your question. About the satellite communication, because the development of Internet in recent years is focus on wireless technology, so satellite network technology is still not popularized yet. And within the computer science field, the technology has not popularized is normally high cost.

We would like to take JSAT Corporation (the biggest satellite communication supply company in the Pacific Ocean area of Asia) for example, which provides a SPACE IP service for enterprise clients. The SPACE IP service is composed of two kinds of plans, light plan and standard plan. Light plan provides 5Mbps download speed bandwidth and 1Mbps upload speed bandwidth, the monthly cost is one hundred thousand yen; standard plan provides 10Mbps of download speed bandwidth and 2Mbps upload speed bandwidth, the monthly cost is two hundred thousand yen. Before starting to use the satellite communication, enterprises also have to pay at least four hundred thousand yen for the construction of equipments. And in addition to the monthly cost for data communication, enterprises have to pay about thirty thousand yen for renting the communication equipments.

Following we explain above, if we compare its light plan with general Internet services (such as DSL or cable modem), for the monthly communication cost, satellite communication is about twenty times of them, but the communication speed is only their 0.05%. Because of the conditions we list above, we claim that satellite communication technology is still not a useful communication way for now.