DEVELOPMENT OF A MARITIME INCIDENT DATABASE USING A RELATIONAL DATABASE MANAGEMENT SYSTEM

Siyu Xu
Department of Transportation, Shipping and Logistics
School of Naval Architecture, Ocean and Civil Engineering
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai, 200240, P.R. China
Tel: 8621-54748155 Fax: 8621-62933163; Email: xusiyu_cathy@163.com

Hao Hu, Corresponding Author
Professor and Head of the Department of Transportation, Shipping and Logistics
School of Naval Architecture, Ocean and Civil Engineering
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai, 200240, P.R. China
Tel: 8621-62933091 Fax: 8621-62933163; Email: hhu@sjtu.edu.cn

Word count: 4,992 words text + 8 tables/figures x 250 words (each) = 6,992 words

Submitted to the Transportation Research Board 96th Annual Meeting for Presentation

Submission date: November 15th, 2016

ABSTRACT

Maritime incidents are still inevitable even though significant progress has been made in the development of maritime technology and management. Consequences of the incidents are often serious and unacceptable. Many kinds of maritime incident databases have been established for people to learn from what has happened and thus develop corresponding mitigation measures. However, our investigation of two widely used databases, i.e., the Global Integrated Shipping Information System (GISIS) and the Lloyd's List Intelligence (LLI), shows that most existing maritime incident databases just record basic information of the incidents in a single table, which may result in data redundancy and impede the extension of the database. This paper develops a relational database management system for maritime incidents (Maritime-RDBMS) that overcomes the identified limitations. The Entity-Relationship (ER) model is firstly used to depict the inter-related semantic information of maritime incidents and a relational database model is subsequently formed. Microsoft Access is employed to implement the proposed database and some functions are designed to demonstrate the application of the database. Our preliminary study shows that the proposed Maritime-RDBMS is implementable and has the potential for practical use.

Keywords: Maritime Incident, Relational Database, ER model, Microsoft Access

1. INTRODUCTION

Characterized by its large capacity and economies of scale, maritime transportation is a primary mode of international transportation and accounts for approximately 90% of global trade volume. Even with the advancement of maritime technology and management, maritime transportation is still recognized as a high-risk mode of transportation. Every year, many incidents take place at sea, often with serious consequences for people, ships or the environment. In fact, people today are less tolerant of maritime incidents than ever before. Maritime safety has become a priority for government authorities, and thus, maritime incident analysis has become a significant area of study. Many efforts have been made by researchers to investigate the causal mechanisms surrounding maritime incidents and the remedial measures that should be applied to prevent these incidents in the future. As compared to experts' judgments and simulation experiments, data from real incident cases are the most important basis for maritime incident analysis performed by researchers.

The International Maritime Organization (IMO) has issued a series of international standards and recommended practices for a safety investigation into maritime casualties in order to collect incident data (*1*). Generally, maritime incident data include primary and secondary data sources (*2*). The primary data are provided by those directly involved in an incident, such as the crew and passengers, or monitored by the equipment installed onboard, such as the Voyage Data Recorder (VDR), the Automatic Identification System (AIS), Vessel Traffic Services (VTS), Very High Frequency communication systems (VHF) and the Electronic Chart Display and Information System (ECDIS). Secondary source is the data processed from the primary ones through multi-information fusion and managed in various databases. These databases underlie maritime incident analysis and they can generally be classified into three types: public databases maintained by the IMO, commercial databases maintained by classification societies and national databases maintained by government agencies (*3*). Among them, the Global Integrated Shipping Information System (GISIS) and the Lloyd's List Intelligence (LLI) are two most well-known databases.

To facilitate direct reporting of incident information by member states, the IMO launched the GISIS in 2005 (*4*). "Marine Casualties and Incidents (MCI)" is one of its twenty modules. In the MCI module, an incident record is composed of four parts: incident summary, reporting forms, investigation reports and analyses. Notably, as defined by the IMO circulars MSC-MEPC.3/Circ.3, reporting forms include ten annexes, focusing on collecting information on different aspects of an incident (*5*). In the GISIS, basic or advanced search based on certain criteria to find a particular type of incidents is allowed and the searching results can be visualized on the map. Unlike the GISIS as a public resource, the LLI, formerly known as the Lloyd's Marine Intelligence Unit (LMIU), is a private commercial database maintained by Lloyd's Register. In this database, the Insurance Channel and the Law & Regulation Channel both include a module called Casualties, which collects all types of maritime incidents with different severities worldwide of merchant ships of over 100 gross tons (*6*). This module contains detailed records on serious casualties and displays the distribution of casualties according to the casualty type, ship type and gross tonnage for last six months. Each incident can be fully described with 33 fields, and click the name of some ship or one of its ownerships, one can reach other interfaces presenting more details concerned. Table 1 summarizes characteristics of the GISIS and the LLI.

TABLE 1 Summary of Characteristics of the GISIS and the LLI

Items	GISIS	LLI
Ways to Collect	Member states directly report incident information to GISIS (Mandatory/ Passively)	Collect from multiple suppliers, like port agents and rescue centers, and combine with terrestrial and satellite AIS tracking (Actively)
Reliability/Accuracy	Average	Good
Comprehensiveness	Average Cover over 9800 records dated from 1973	Good Cover over 86400 records dated from 1965
Update Frequency	Average	Good
Availability	Public	Private
Recording Principle	Record incidents from the perspective of the incident	Record incidents from the perspective of the ship
Definition of Seriousness	Four categories: -Very serious casualties -Serious casualties -Less serious casualties -Marine incidents	Two categories: "Seriousness" involves only what may affect the ship, and not the ship's crew or passengers
Causes of Incidents	Analyze from multiple aspects, including people, ships, cargoes and the environment	Record at most 3 causes, which actually are confused with casualty type
Consequences of Incidents	Qualitative: Consequences to ships Quantitative: Consequences to people and the environment	Qualitative: Fatality Indicator, Pollution Indicator and Serious Indicator (Ship) Quantitative: Consequences to people

In common with other structured databases, the GISIS and the LLI both report maritime casualties by a set of data fields, covering worldwide incidents. In both databases, data fields concerning the incident and the ship are encompassed in two separate tables, which are correlated by the ship's unique identifier, i.e. the IMO number or the ship name. This reveals that the GISIS and the LLI both have applied some principles of relational database technology. Nevertheless, the GISIS and the LLI might only take advantage of preliminary knowledge about the relational database. The GISIS records casualties from the perspective of the incident but cannot expose the incident history of each ship intuitively. Contrastingly, the LLI lists the incident history of each ship but fail to uncover other possible ships involved in the same incident. These problems would be solved if the GISIS and the LLI were developed systematically as relational databases.

Most maritime incident databases may utilize the basic ideas of a relational database, but they fail to do so in any systematic way. First, existing databases were not designed systematically and thus fail to present a scientific database model. Information surrounding an incident is usually recorded in a single table, which may result in data redundancy and impede the extension of the database. In addition, tables recording different incidents are isolated from one another, making it difficult to form a network of information to be used in knowledge management. Furthermore, most incident databases are merely used to support researchers who wish to retrieve incident records for further study based on certain criteria; however, these databases lack the type of flexible analysis functions required by an intelligent database. Considering that databases will be updated over time, the traditional method that exporting incident data from the databases and analyzing them externally makes it difficult to achieve real-time analytical results. Therefore, to best make use of the latest data, it is necessary to be able to conduct analysis automatically in the database itself.

The purpose of this paper is to summarize the properties and capabilities of the existing maritime incident databases and to establish an improved relational maritime incident database called the Maritime-RDBMS using a systematic approach. The Entity-Relationship (ER) model is employed to depict the inter-related semantic information surrounding maritime incidents and to provide a concise visualization of concerned entities and their relationships. Full use of relational database technology is employed to accomplish the database design. With this design, incident information can be organized in a logical way that strengthens data independence while maintaining data correlation. The preliminary study of the proposed database finds that such a design can be used to reflect the real world more objectively and scientifically and that such a design makes it possible to implement an intelligent database for use in knowledge management.

The remainder of the paper is organized as follows: Section 2 reviews the literatures related to the development of maritime incident databases and its typical applications as well as limitations in existing databases. Next, the ER model of the proposed relational maritime incident database is developed and converted to the relational database model in Section 3, and Section 4 implements the database design through the software Microsoft Access and imports empirical data to demonstrate some sample applications. Finally, conclusions and recommendations for future work are provided in Section 5.

2. LITERATURE REVIEW

Incident databases are the important basis of maritime incident analysis and the goal can be better achieved by effect of scale of databases. In addition to the databases on the global level, e.g., the GISIS and the LLI (*6*), more databases specific to some country or region are available. Maritime authorities, such as National Transportation Safety Board in the U.S. and Marine Accident Investigation Branch in the U.K., investigate maritime incidents involving ships under their flag worldwide and all ships in their territorial waters and issue investigation reports that compose the national maritime incident database. Other incident databases include the Finnish DAMA database, the Helsinki Commission (HELCOM) Database that gathers statistics of Baltic Sea incidents, the European Marine Casualty Information Platform (EMCIP) database and so on (*7*). Moreover, existing maritime incident databases are constantly improved, for example, IMO made some amendments to MSC-MEPC.3/Circ.3 to revise the online reporting procedures of the GISIS in 2013 (*8*). Some advanced information technologies, such as the Geographic Information System (GIS), are also integrated into maritime incident databases to develop management information systems for safer maritime transportation (*9*, *10*).

Applications of maritime incident databases are multifaceted, typically focusing on data processing, statistics and risk analysis surrounding maritime incidents. When people utilize incident databases, preliminary data processing such as data fusion or text mining is often required. Li *et al.* applied the Dempster-Shafer theory (DST), a generalization of the Bayesian theory, to combine evidence from different databases for a comprehensive result (*11*). Text-mining techniques were also utilized to replace human efforts in extracting the key information from investigation reports which are normally in text format (*12-14*). Apart from data processing, statistics methods are often used to summarize risk factors and reveal the causal mechanism. Zhang *et al.* explored ship incident frequency according to different incident types and consequences. And a series of statistic examinations were conducted to figure out relationships between contributory factors and incident consequences (*2*). Based on incident databases, risk profile can also be realized. As for risk of ships, Li *et al.* proposed a quantitative safety index for each worldwide ship using the binary logistic regression method (*15*). As for risk of ocean regions, Huang *et al.* visualized the spatial distribution of worldwide maritime incidents in the GISIS between 2002 and 2011 and carried out hot spot analysis and buffer analysis on the GIS platform (*3*).

Although constantly improved, maritime incident databases still suffer from some problems, such as underreporting of incidents (*6*, *16-17*), the lack of a unified recording standard, etc. In the studies, common methods used to estimate the actual number of incidents include the conditional probability method, the capture-recapture method, the best case scenario method and the up-scaling of subset data method. It is estimated that approximately 50% of all occurred incidents are underreported, and users of incident databases are suggested to assume a certain degree of underreporting to adjust their analyses (*17*). As for the recording standard of incident data, different incident databases may suffer from a lack of the uniformity in recording methods and field taxonomies (*7*). Far more importantly, analysis results will be handicapped if data are incomplete. Devanney outlined the problems with ship casualty data, e.g., the censored data, unauditable data and confusing cause and effect, and a public database of tanker casualties was developed as an example to demonstrate what a reasonable database should be like (*18*). Also, Su *et al.* developed a management information system of maritime incidents based on SQL Server 6.0 and Delphi 6.0 (*19*). But the system only provided a brief function architecture used for basic information management without following a systematic methodology of design to facilitate data retrieval and analysis. . The standardization of incident reporting with respect to road transportation also has a certain relevance to maritime incident reporting. To standardize traffic crash location records, a five-element crash location description method based on the linear referencing system was defined and taught to practitioners for use in field operations (*20*). Additionally, unifying the field naming and the definition of attributes of the databases in the transportation industry was also studied (*21*).

In general, as the availability and transparency of incident data are improved (22), many types of maritime incident databases are becoming available and greatly stimulating evidence-based research. Whereas the deficiencies of databases are likely to negatively affect the results of applications that use them, great efforts have been made to improve maritime incident databases from many different perspectives. Almost all the existing academic studies focus on eliminating the underreporting of incidents, standardizing recording methods, etc., but many overlook the necessity of improving the storage structure of incident information. Actually, creating a reasonably structured database model based on relational database technology can make maritime incident databases fundamentally more practical.

3. DEVELOPMENT OF THE RELATIONAL MARITIME INCIDENT DATABASE

The relational database, first defined by Codd in 1970 (*23*), has been the predominant type of database. It is based on the relational model, whose central data description construct is the relation, i.e., the table. Data are stored in multiple correlative tables, where records are distinguished from each other through the primary key and relate to the corresponding records in other tables through the foreign keys. The relational database has good data independence and low data redundancy, and manages data in a systematic and interrelated manner. A relational database design normally includes six steps: requirements analysis, conceptual database design, logical database design, physical database design, implementation and operational maintenance. In particular, the conceptual and logical database designs incorporated in this section are critical to the overall design of a system.

3.1 Conceptual Database Design with the ER Model

3.1.1 The ER Diagram

Unlike existing databases, the proposed relational maritime incident database aims not only to record maritime incident information comprehensively but also to improve data independence and logic, while maintaining the correlation among data, so as to facilitate data query and statistics. This objective will guide the whole design process. Conceptual database design is the initial design phase, in which a well-known semantic model called the ER model is used to describe and abstract the data in the real world (*24*). As shown in Figure 1, the ER model using the Chen notation pictorially presents the entities relevant to a maritime incident and how they relate to one another. According to Chen's diagramming technique, the ER diagram consists of entities, relationships and attributes, which are represented by rectangles, diamonds and ovals, respectively.

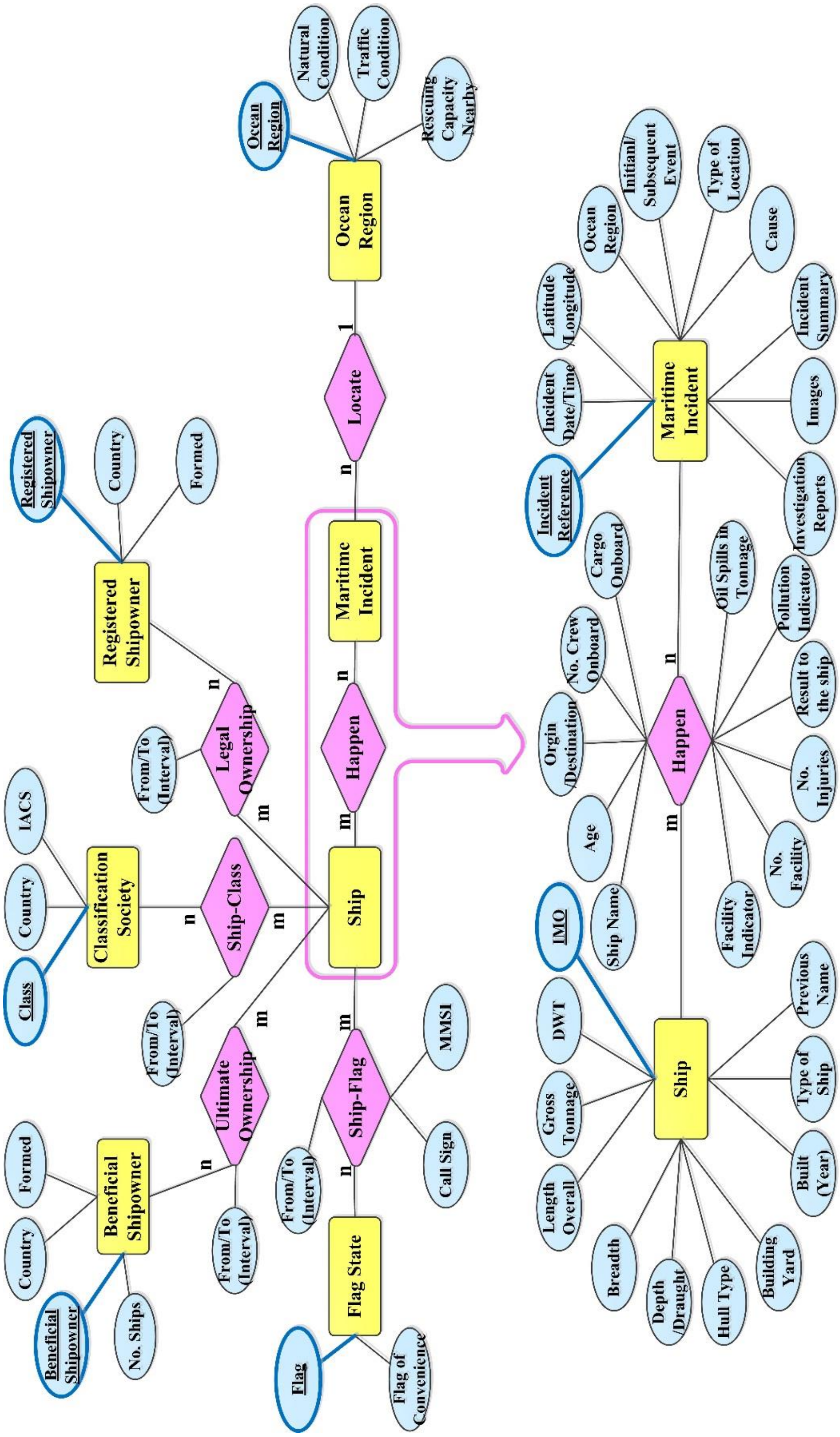


FIGURE 1 The ER diagram of the relational maritime incident database.

3.1.2 Entities in the ER Diagram

An entity is an object in the real world that is distinguishable from other objects, either a physical object or a concept (*25*). In the context of a maritime incident, the maritime incident and the ships involved are the most important entities, and they are emphasized with a box and analyzed in more detail at the bottom of the diagram.

One certainly cannot claim that the flag is a possible cause of a maritime incident, but it is one quite important risk factor. The flag may be chosen as a "proxy" for other variables that cannot be easily measured, such as crew composition, crew training and others. Similarly, the classification society and the country of ownership of a ship may also be regarded as "proxy" variables and they have proven to be statistically significant factors affecting maritime incident frequencies (*26*, *27*). Therefore, the flag state, classification society, registered ownership and beneficial ownership, i.e., the real ownership, are considered as the primary entities most closely related to the ship in terms of maritime safety.

Additionally, the ocean region is identified as the major entity related to the maritime incident, because maritime administrations are quite concerned with the condition of incident sites, such as natural conditions and traffic conditions. In accordance with the *World Casualty Statistics* by Lloyd's Register, worldwide ocean region is divided into 31 zones. Due to their unique geographical environment, different region will have different effects on a ship's safety (*28*). Similarly, the LLI classifies the location of casualties into 34 regions, and this taxonomy will be employed in this paper so as to associate maritime incidents with spatial information that can be used in further analysis.

3.1.3 Relationships in the ER Diagram

With the relevant entities identified, the next step is to recognize the logical connection among these entities, i.e., the relationships. The relationship that an entity has with another entity is usually realized by the primary key and the foreign key. When a primary key migrates to another table, it becomes a foreign key in the other table. With respect to a maritime incident, it is not difficult to determine relationships among entities; however, a heavy emphasis is also placed on determining the cardinality of such relationships. Since the relationships in the established ER model are all binary, they can be referred to as being one-to-one (1:1), one-to-many (1:n) or many-to-many (m:n) and labelled in the ER diagram.

A ship may change its flag, classification and/or ownership status several times during its lifetime, and any given ship may also be involved in more than one incident, hence the relationships between the ship and other connected entities all belong to the many-to-many relationship. Additionally, any casualty can be located to a region, while any region may have seen more than one casualty, so the relationship between the ocean region and the incident can be defined as a one-to-many relationship.

3.1.4 Attributes in the ER Diagram

Apart from entities and relationships, another essential element of an ER diagram is the attribute used to describe entities and relationships. A comprehensive set of attributes is included to describe the maritime incident, the ship involved and their relationship for the reason that the choice of attributes reflects the level of detail at which one want to represent information about entities and relationships. These attributes are designed on the basis of the recording format of several widely used maritime incident databases, especially the GISIS and the LLI. Modifications are also made to improve the recording pattern of some data fields, particularly the incident causes and consequences. When identifying attributes, the attribute domain is simultaneously specified for each attribute, including the allowable set of values, the size and the format.

As modern maritime technology tends to provide more safeguards against possible errors, in which case the occurrence of an incident requires an even larger combination of factors (*29*). Hence, a systematic record of contributing factors of casualties is necessary. From the perspective of man-machine-environment systems engineering, an improved taxonomy designed for collecting causal factors is presented in Figure 2 (*30*). Also, a systematic method is employed to record incident consequences, taking into account the consequences to people, ships and the environment. Without defining the seriousness of incidents, this paper only uses the “facility indicator”, “pollution indicator” and “result to the ship” (i.e., total loss/unfit to proceed/remain fit to proceed) to qualitatively indicate the incident results for the convenience of data retrieval. And details about consequences to people and the environment are also collected quantitatively. Moreover, since data fields can hardly provide a detailed context for each incident, the Maritime-RDBMS also allows users to attach investigation reports and incident images to enhance the amount of contextual information available about any given incident.

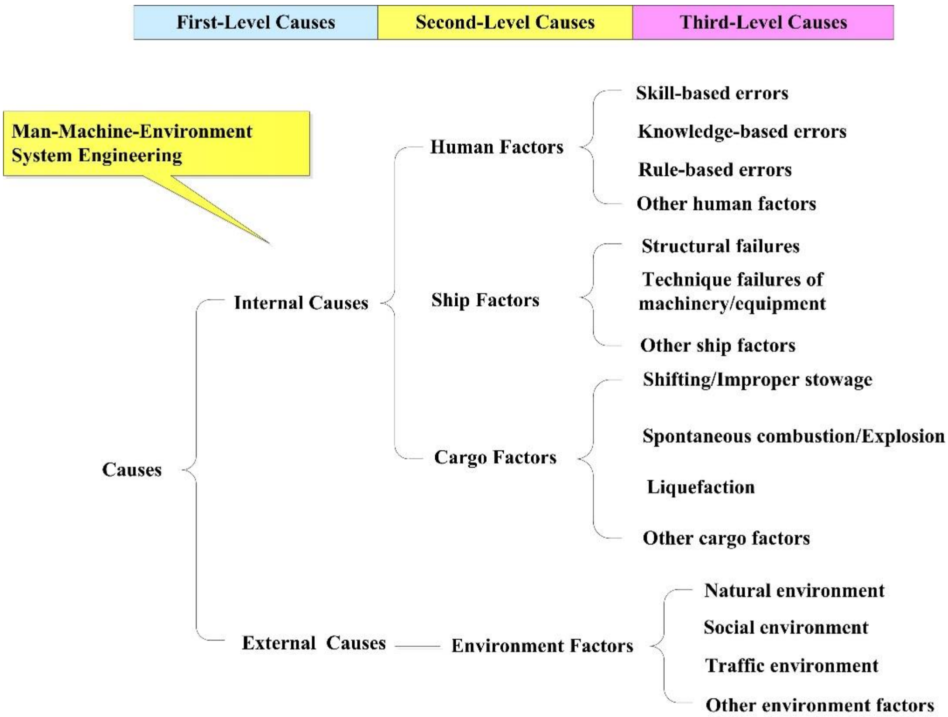


FIGURE 2 Alternative contributing factors of a maritime incident.

3.2 Logical Database Design: ER Model to Relational Model

3.2.1 The relation schema

When designing a relational database, the objective of the logical database design is to map a conceptual data model onto a relational data model, i.e., to translate the ER model into relation schemas. The relation schema describes the column headers for the table, which is actually the recording format of maritime incidents. A relation schema can be expressed as: name of the relation (attribute 1, attribute 2, ..., attribute n). In accordance with the ER model of the developed maritime incident database, relation schemas are derived as follows:

(1) Entities:

- Ship (**IMO**, Dead Weight Tonnage, Gross Tonnage, Length Overall, Breadth, Depth, Draught, Hull Type, Building Yard, Built (Year), Type of Ship, Previous Name)
- Maritime Incident (**Incident Reference**, Incident Date, Incident Time, Latitude, Longitude, Ocean Region, Type of Location, Initial Event, Subsequent Event, Cause, Incident Summary, Images, Investigation Reports)
- Flag State (**Flag State**, Registry Policy)
- Classification Society (**Class**, Country, International Association of Classification Societies (IACS))
- Beneficial Shipowner (**Beneficial Shipowner**, Country, Formed (Year), Number of Ships)
- Registered Shipowner (**Registered Shipowner**, Country, Formed (Year))
- Ocean Region(**Ocean Region**, Natural Condition, Traffic Condition, Rescuing Capacity Nearby)

(2) Many-to-many relationships:

- Happen (**IMO**, **Incident Reference**, Ship Name, Age, Origin, Destination, Number of Crew Onboard, Particulars of Cargo Onboard, Facility Indicator, Number of Fatalities (Dead or Missing), Number of Injuries, Result to the Ship, Pollution Indicator, Oil Spills in Tonnage)
- Ship-Flag (**Flag State**, **IMO**, From, To, Call Sign, MMSI)
- Ship-Class (**Class**, **IMO**, From, To)
- Ultimate Ownership (**Beneficial Shipowner**, **IMO**, From, To)
- Legal Ownership (**Registered Shipowner**, **IMO**, From, To)

An entity can be mapped to a relation in a straightforward way. Each attribute of the entity becomes an attribute of the table, and the primary key is underlined. With respect to relationships, a typical way to deal with the one-to-many relationship is to insert a foreign key into the table that represents the “many” side of the relationship. And a many-to-many relationship needs to be transformed into two one-to-many relationships, which can be achieved by creating an additional relation. The attributes of such relation contain the primary keys of participating entities and the descriptive attributes of the relationship.

3.2.2 Normalization and Integrity Constraints

After transformation, relation schemas are checked against normalization and integrity constraints to make sure all the relations are correct structurally. Firstly, attributes in each relation are validated using the rules of normalization to eliminate non-atomic values and data redundancy. Functional dependencies and the primary keys of each relation are used in the process of normalization. To achieve a balance between minimal data redundancy and maximum accessing efficiency, each relation conforms to the rules of third normal form, 3NF.

In addition, in order to prevent the database from becoming incorrect, invalid or inconsistent, the relational data model is examined against integrity constraints, including entity integrity, referential integrity and user-defined integrity. In particular, as entity integrity specifies, the primary key must not be an empty set of attributes. Assuming that ships without an IMO number replace this field with the ship name, the IMO number can be regarded as the primary key of the ship relation.

4. APPLICATION OF THE DATABASE DESIGN

After completion of the basic design of the relational maritime incident database Maritime-RDBMS, this section will implement the database design and conduct a few sample applications to demonstrate that the developed database is implementable and has the potential for practical use.

4.1 Implementation of the Basic Database Design

With the basic database design in hand, Microsoft Access, a well-known Relational Database Management System (RDBMS), is utilized to implement the design and to implement and execute sample applications that query the database. An Access database contains six types of objects: tables, queries, forms, reports, macros and modules. Tables are the ultimate data structures in which data are stored. The structures of the database tables are specified according to the previously designed relation schemas. These tables are then correlated through the primary and foreign keys. The relationships between pairs of tables created in Access are shown in Figure 3.

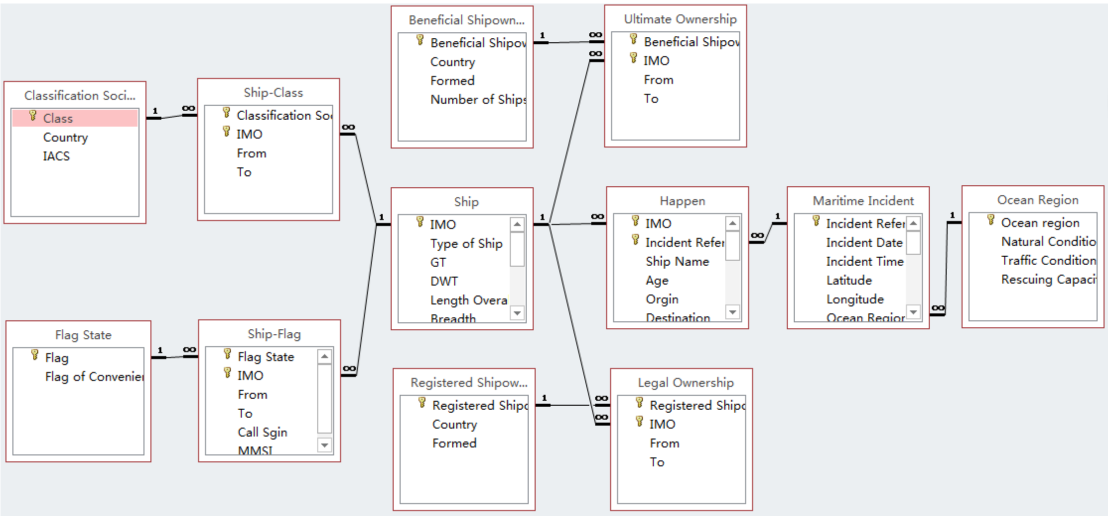


FIGURE 3 Establishment of tables and relationships in Microsoft Access.

4.2 Establishment of Queries and Forms for the Database Application

Apart from tables, queries and forms are also utilized to build the target database and to create some practical applications. Using the aforementioned objects, the proposed database launches two modules, an information management module and a statistical analysis module, as shown in Figure 4. Furthermore, all the serious incidents with oil spills in the LLI database that occurred from January 1st, 2014 to December 31st, 2015, a total of 60 incidents, are imported into the Maritime-RDBMS to be used for demonstration.

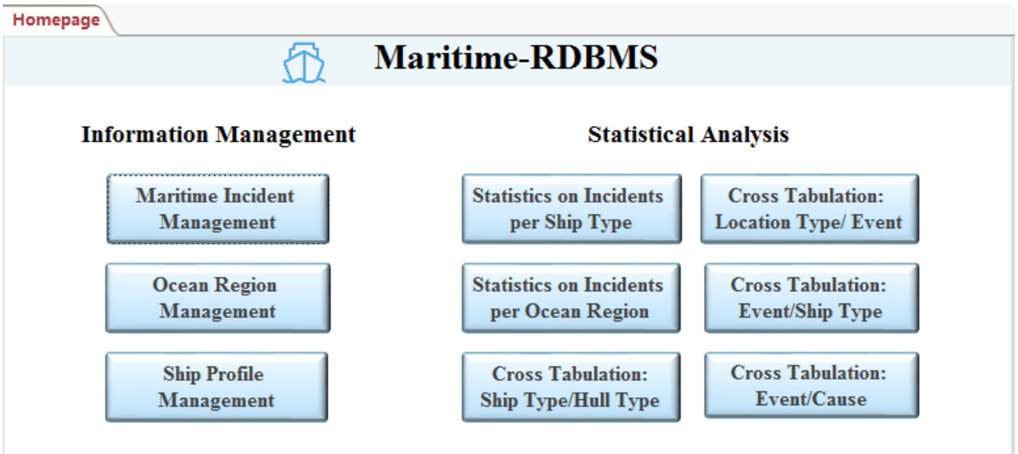
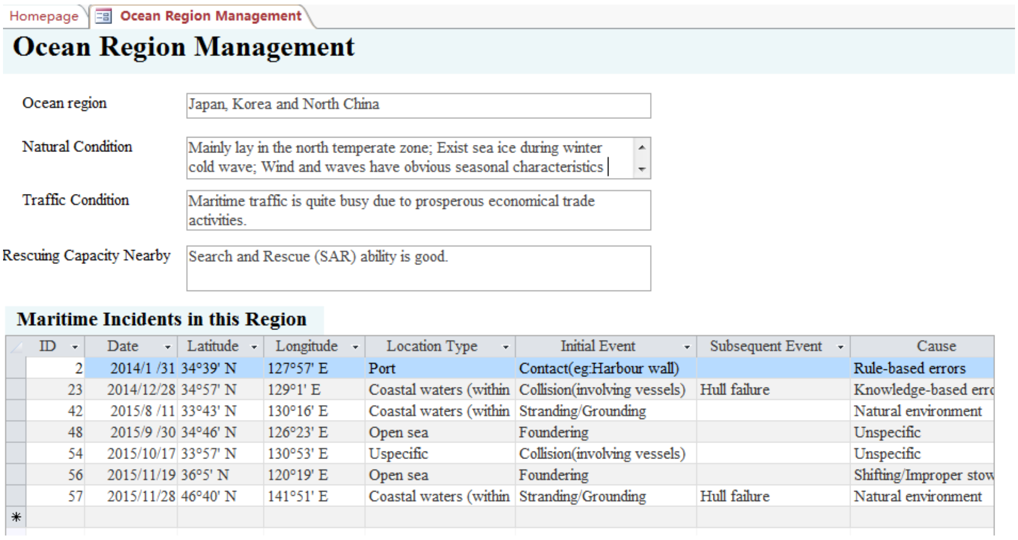


FIGURE 4 Homepage of the database application.

The information management module includes three functions. The first function, Maritime Incident Management presents comprehensive incident data as well as basic profiles of the ships involved in a given incident. Notably, for a ship whose information has been stored in the system, the administrator simply needs to input the IMO number of the ship in order for the system to automatically match this number with the data of the ship that has already been stored in the database. Such functionality largely relieves the user of the burden of data entry. Compared to the first function, the two other functions offered by the database application are more innovative, providing the ability to manage incidents from multiple perspectives. As illustrated in Figure 5, Ocean Region Management can provide both the basic information of the region and corresponding incident records in each region. This function is helpful for coastal states to improve the safety level of the ocean region nearby, and ships passing by can learn about potential risk from the past incidents and raise their vigilance accordingly. Another function, Ship Profile Management, is designed to document the flag, class and ownerships of each ship during different periods as well as the ship's incident history. This function may serve as a reference in estimating the risk level of a given ship.



ID	Date	Latitude	Longitude	Location Type	Initial Event	Subsequent Event	Cause
2	2014/1 /31	34°39' N	127°57' E	Port	Contact(eg:Harbour wall)		Rule-based errors
23	2014/12/28	34°57' N	129°1' E	Coastal waters (within	Collision(involving vessels)	Hull failure	Knowledge-based erro
42	2015/8 /11	33°43' N	130°16' E	Coastal waters (within	Stranding/Grounding		Natural environment
48	2015/9 /30	34°46' N	126°23' E	Open sea	Foundering		Unspecific
54	2015/10/17	33°57' N	130°53' E	Uspecific	Collision(involving vessels)		Unspecific
56	2015/11/19	36°5' N	120°19' E	Open sea	Foundering		Shifting/Improper stow
57	2015/11/28	46°40' N	141°51' E	Coastal waters (within	Stranding/Grounding	Hull failure	Natural environment
*							

FIGURE 5 Interface of the Ocean Region Management function.

The statistical analysis module provides two aggregate queries and four crosstab queries as sample applications, which can be programmed with the Structured Query Language (SQL). Aggregate queries count the incidents grouped by a specific field, e.g., the ship type (Figure 6) or the ocean region (Figure 7), and perform descriptive statistics on the results. The function "Statistics on Incidents per Ship Type" counts the incidents during a specified period according to the type of ship and reports the age distribution of the ships involved in such incidents and the cumulative incident consequences. The function "Statistics on Incidents per Ocean Region" counts the incidents and ships involved according to the ocean region. Moreover, it considers the seriousness of incidents by counting the number of ships involved in incidents that produced fatalities or in incidents where the entire ship was lost and calculates the ratio of the number of ships involved in each of these types of incidents to the overall number of ships involved in all types of incidents in the region. In this case, one would be able to identify the "hot spots" more reasonably for the reason that if two regions have similar quantity of incidents, but one reflects more serious consequences, this region is more likely to be a "hot spot". In addition to the two aforementioned aggregate queries, the system also provides four crosstab queries, which are targeted at finding out some regularities between two kinds of data fields and deriving some design modifications and administrative suggestions.

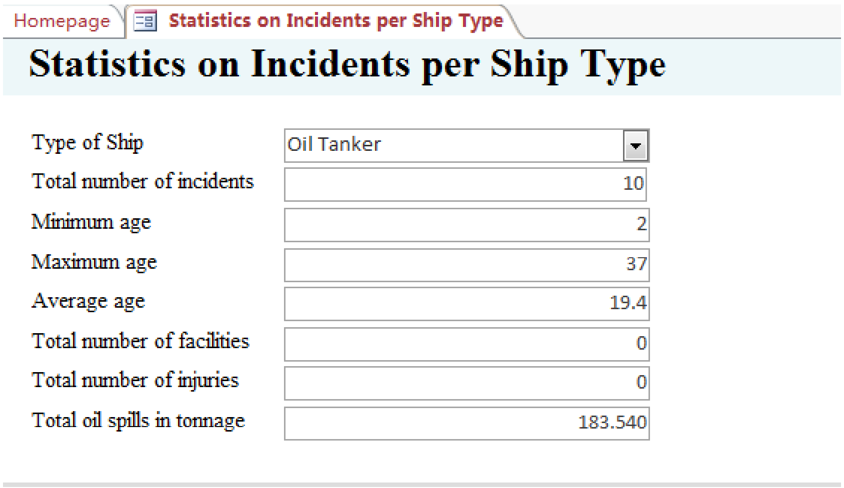


FIGURE 6 Interface of Statistics on Incidents per Ship Type function.

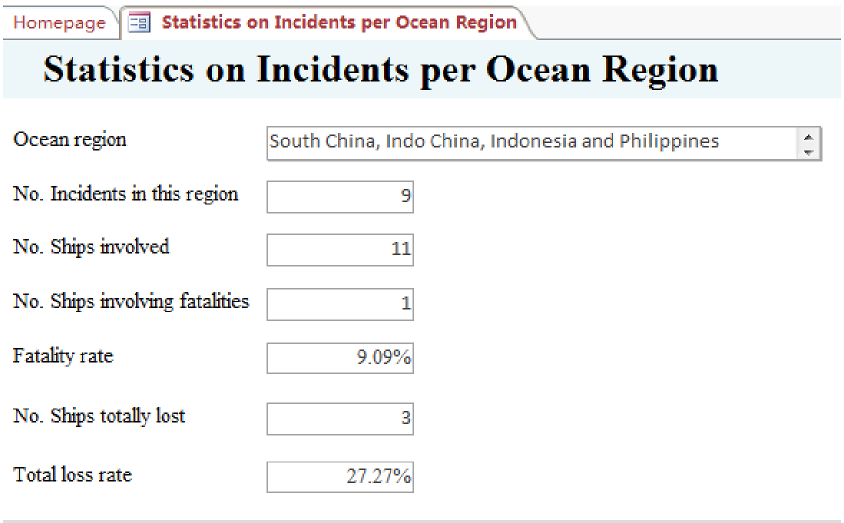


FIGURE 7 Interface of Statistics on Incidents per Ocean Region function.

Compared to existing maritime incident databases, this improved Maritime-RDBMS has two advantages: (1) Thanks to the characteristics of the relational database, information regarding different aspects of an incident can be stored separately and yet still remain correlated through common fields shared by the database. Such a design creates a network of information that allows one to implement better information management functions while still maintaining data independence. (2) Unlike other maritime incident databases that use tables as the ultimate data source, all the queries provided by our database application generate dynamic data sets. Thus, the improved database can provide real-time statistical functions that are not available in existing databases. Once incident data are updated, various queries will present the latest statistical results. Such an approach is more convenient and responsive than much evidence-based research performed outside the incident database itself.

5. CONCLUDING REMARKS

This paper developed a relational maritime incident database called the Maritime-RDBMS with the application of a relational database management system and provided some database applications to interact with the database. Three important conclusions are drawn as follows.

- Generally, three types of maritime incident databases exist: public databases, commercial databases and national databases. This paper makes an in-depth exploration of two well-known databases, the GISIS and the LLI, and highlights some limitations that exist in most databases. By analyzing the data semantics of maritime incidents in a structured way, an improved relational maritime incident database is developed to address the shortcomings of the aforementioned systems. In addition, some useful database applications are designed as examples that demonstrate the utility of the proposed database.
- The relational database is the predominant type of database in use today. Although most maritime incident databases have reflected the basic idea of the relational database, they fail to apply this technology systematically. In this research, the ER model is applied to help understand the data semantics of maritime incidents in a structured way. A standard relational database Maritime-RDBMS is then developed to record maritime incidents in a logical way, which can help reduce data redundancy in maritime incident recording and facilitate the extension of the database. Moreover, the developed database may be used not only to collect incident information but also to conduct knowledge management in an intelligent way.
- The MCI module in the GISIS is widely used in maritime safety management and academic studies. Established in 2005, the incident reporting format of this module has been constantly improved, and its data fields are well defined. However, few efforts have been made to improve the storage structure of the data stored in this database or the statistical functions provided by this database. Therefore, it is suggested that IMO follow the pattern of the relational maritime incident database design proposed in this research to further improve the MCI module in the GISIS.

This paper demonstrates the potential for systematically designing relational database models to be used in maritime incident databases. Only 60 incident records were imported in this research for demonstration purposes without generating statistically significant conclusions. Further work about maritime incident analysis using the proposed database should be performed in the future. Additionally, converting the proposed database design to a web database with a scripting language is also necessary to support online operations for database users.

REFERENCES

1. IMO. Adoption of the Code of the International Standards and Recommended Practices for a Safety Investigation into a Marine Casualty or Marine Incident (Casualty Investigation Code). Resolution MSC. 255 (84), 16 May. 2008, pp. 1-24.
2. Zhang, L., H. Wang, and Q. Meng. Evidence-Based Ship Accident Consequence Analysis Using Heterogeneous-Source Accident Reports (#16-3091). *Proceedings of the 95th Annual Meeting of Transportation Research Board*, Washington D.C., 2016.
3. Huang, D. Z., H. Hu, and Y. Z. Li. Spatial Analysis of Maritime Accidents Using the Geographic Information System. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2326*, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 39-44.
4. IMO. Global Integrated Shipping Information System (GISIS). Resolution A. 1029 (26), 26 Nov. 2009, pp. 1-2.
5. IMO. Marine Casualties and Incidents. https://gisis.imo.org/Public/MCI/Default.aspx. Accessed on 10 July, 2016.
6. Qu, X. B., Q. Meng, and S. Y. Li. Analyses and Implications of Accidents in Singapore Strait. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2273*, Transportation Research Board of the National Academies, Washington D.C., 2012, pp. 106-111.
7. Ladan, M. and M. Hänninen. Data Sources for Quantitative Marine Traffic Accident Modeling. 2012.
8. IMO, Reports on Marine Casualties and Incidents. MSC-MEPC.3/Circ.4/Rev.1, 18 Nov. 2014, pp. 1-31.
9. Dobbins, J. P., and M. D. Abkowitz. Development of an Inland Marine Transportation Risk Management Information System. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1782, Transportation Research Board of the National Academies, Washington D.C., 2002, pp. 31-39.
10. Martin, P. H., E. J. LeBoeuf, E. B. Daniel, J. P. Dobbins, and M. D. Abkowitz. Development of a GIS-based Spill Management Information System. *Journal of hazardous materials*, Vol. 112, No. 3, 2004, pp. 239-252.
11. Li, Y. Z., H. Hu, and D. Z. Huang. Data Fusion of Maritime Incident Databases with Dempster-Shafer Theory. In *Transportation Research Record: Journal of the Transportation Research Board*, No.2426, Transportation Research Board of the National Academies, Washington D.C., 2014, pp. 27-32.
12. Mazaheri, A., J. Montewka, J. Nisula, P. Kuhala. Usability of Accident and Incident Reports for Evidence-based Risk Modeling–A Case Study on Ship Grounding Reports. *Safety science*, Vol.76, 2015, pp. 202-214.
13. Tirunagari, S., M. Hanninen, K. Stanhlberg, and P. Kujala. Mining Causal Relations and Concepts in Maritime Accidents Investigation Reports. *International Journal of Innovative Research and Development*, Vol.1, No. 10, 2012, pp. 548-566.
14. Tirunagari, S., M. Hanninen, A. Guggilla, K. Stahlberg, and P. Kujala. Impact of Similarity Measures on Causal Relation Based Feature Selection Method For

Clustering Maritime Accident Reports. *Journal of Global Research in Computer Science*, Vol.3, No. 8, 2012, pp. 46-50.

15. Li, K. X., J. B. Yin, and L. X. Fan. Ship safety index. *Transportation research part A: policy and practice*, Vol. 66, 2014, pp. 75-87.
16. Psarros, G., R. Skjong, and M. S. Eide. Underreporting of Maritime Accidents. *Accident Analysis and Prevention*, Vol. 42, No. 2, 2010, pp. 619-625.
17. Hassel, M., B. E. Asbjørnslett, and L. P. Hole. Underreporting of Maritime Accidents to Vessel Accident Databases. *Accident Analysis and Prevention*, Vol. 45, No. 6, 2011, pp. 2053-2063.
18. Devanney, J. Uses and Abuses of Ship Casualty Data. *Center for Tankship Excellence*, July. Available at www. c4tx. org, 2008.
19. Sun, H. N., X. F. Duan, and Z. J. Liu. SQL-based Management System and Implementation of Maritime Accident. *Journal of Dalian Maritime University*. Vol.33, 2007, pp. 120-121.
20. Zhang, B., X. S. Wang, Y. Luo, X. H. Chen, J. Chen, and S. R. Wang. Method of Urban Traffic Crash Location Record Description. *China Safety Science Journal*, Vol.22, No. 2, 2012, pp. 127-133.
21. Wang Y. Y., H. X. Zhao, S. Y. Zhang, and W. L. Song. Specification for Field Naming and Attribute Definition of Database for Traffic & Transportation Information System. *Transport Research*, Vol.1, No. 3, 2015, pp. 89-96.
22. Wan, Z. China's Scientific Progress Hinges on Access to Data. *Nature*, Vol. 520, No. 7549, 2015, pp. 587
23. Codd, E. F. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, Vol.13, No. 6, 1970, pp. 377-387.
24. Chen, P. P. S. The Entity-Relationship Model—Toward a Unified View of Data. *ACM Transactions on Database Systems (TODS)*, Vol.1, No. 1, 1976, pp. 9-36.
25. Ramakrishnan, R., and J. Gehrke. *Database management systems*. McGraw-Hill, 2000.
26. Li, K. X., J. B. Yin, H. S. Bang, Z. L. Yang, and J. Wang. Bayesian Network with Quantitative Input for Maritime Risk Analysis. *Transportmetrica A: Transport Science*, Vol. 10, No. 2, 2014, pp. 89-118.
27. Psaraftis, H. N., G. Panagakos, N. Desypris, and N. Ventikos. An Analysis of Maritime Transportation Risk Factors. *The Eighth International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 1998.
28. Yin, J. B. Quantitative Risk Assessment for Maritime Safety Management. Lap Lambert Academic Publishing. 2013.
29. Psaraftis, H. N., P. Caridis, N. Desypris, G. Panagakos, and N. Ventikos. The Human Element as a Factor in Marine Accidents. In *Proceedings of the 1st Greek Maritime Transport Conference*. 1998.
30. Rasmussen, J. Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 3, No. 3, 1983, pp. 257-266.