Research on the Construction of Ocean Big Data Platform and Disaster Prevention and Mitigation Application Based on Internet of Things

Chen Chen*, Yao Zhang, Xiaoyong Shi
National Marine Hazard Mitigation Service, Beijing 100194, China

*Corresponding Author: chenchen@nmhms.org.cn

Abstract. The marine disaster prevention and mitigation information service system is a system that can provide information services for marine disaster management decision-making departments and the people. The paper further puts forward the capability objectives of the digital ocean system based on the big data platform of the Internet of Things and the overall framework of the six-level system including the sensing layer, transmission layer, data layer, and application layer, and analyses the safety of the big data platform based on the ocean of things Risks put forward the development requirements for the security of marine big data platforms, and provide references for the planning and construction of marine security systems in the big data era.

Keywords: Internet of Things, ocean big data, big data platform, disaster prevention and mitigation, information service system.

1. Introduction
my country has a vast territory, of which the sea area is 2.997 million km², which is equivalent to one-third of the land area. Strengthening the management of the sea area and effectively preventing and controlling marine disasters is of positive significance for ensuring national safety and maintaining social harmony and stability. In recent years, marine disasters have occurred frequently. According to survey statistics, in 2016, my country’s economic losses caused by marine disasters were as high as 5 billion yuan, and in these disasters, the number of dead or missing reached 60. Based on this situation, it is imminent to start research on the marine disaster prevention and mitigation information service system and step up the construction of a regional disaster prevention system. At present, my country has conducted surveys on the needs of marine disaster prevention and mitigation in coastal provinces and cities, and has achieved certain results [1]. The theory and methods of the construction of the marine disaster prevention and mitigation information service system are gradually improved.

2. Ocean Big Data Theory
The ocean occupies about 71% of the total surface area. In order to understand, understand and use the ocean, humans use satellite remote sensing, aviation, meteorological balloons, stations, buoys, ships, underwater exploration and other methods to observe and investigate the ocean. A very large ocean observation system has been accumulated, and a large amount of ocean scientific data has been accumulated, including field observation data, ocean remote sensing data, numerical model forecast
results and products. These ocean data have the characteristics of massive, multi-source, polymorphic, heterogeneous, etc. With the development and progress of science and technology, people have realized that the ocean is critical to climate prediction and is a key area for understanding global changes. With the continuous improvement of ocean observation accuracy, the rapid shortening of time intervals, and the rapid growth of ocean observation network nodes, the amount of ocean data is growing faster than the growth of big data in other industries. Ocean data has the "5V" characteristics of big data. Ocean big data is the practice and value realization of big data in the ocean field under the theoretical guidance and technical support of big data.

With the development of information technology, countries are racing to launch marine satellites and multiple resource and meteorological satellites with ocean observation functions. They use sensors to conduct three-dimensional detection of the sea surface, extract environmental information such as ocean water colour and ocean dynamics, and obtain ocean space and time data. Satellite remote sensing has the unmatched advantages of other traditional ocean field surveys and other methods. It can conduct large-scale and long-term observations of the global ocean, providing an irreplaceable data source for humans to deeply understand and understand the ocean. In recent years, satellite remote sensing technology has developed rapidly, and the accompanying ocean data scale has exploded [2]. The application value in the fields of ocean environment forecasting, ocean operation and production, and ocean disaster prevention and mitigation has also become increasingly prominent and prominent. Ocean remote sensing is one of the most important data sources of current ocean big data, and it is also an important way to realize digital ocean construction and ocean informatization.

3. Establishment of a big ocean data platform based on the Internet of Things
How to use big data-related technologies and combine the characteristics of ocean applications to quickly and timely analyse and process massive, multi-source, heterogeneous ocean observation and simulation data, realize the development and use of ocean big data, and improve the timeliness of ocean forecasts and applications Sex and accuracy are the key issues in the construction of marine big data platforms. Based on the concept of big data, it is first necessary to integrate the massive and complex multi-source heterogeneous ocean big data, adopt the big data distributed storage + cloud computing platform model, and establish the ocean big data sharing mechanism framework; On this basis, it realizes the comprehensive query and analysis of multiple types of marine information, explores the potential value of marine big data, and provides application services for relevant units, groups and individuals in the marine field, including marine information consulting, marine resource development, marine monitoring and early warning, and marine rights protection Command, marine disaster prevention and mitigation, etc.

3.1. Ocean Big Data Platform Architecture
In general, the hardware and software platform architecture of ocean big data can be divided into three levels: data layer, technical layer and application layer (see Figure 1). The data layer is the basis of the ocean big data platform, including data from various ocean data collection platforms, such as ocean remote sensing data, physical data, and biological data observed by ground, sky, air, coast, ship, oil and gas platform, underwater movement, etc. Chemical data, etc.; through data pre-processing technology, the multi-source heterogeneous ocean big data is organized and managed in a unified mode. The technical layer includes technologies such as the fusion, analysis, mining, forecasting, and visualization of multi-source heterogeneous ocean big data, as well as the integration of related technologies and cloud platform development, to achieve personalized retrieval of ocean data, and to accurately forecast ocean elements [3]. The application layer is based on data retrieval and technology integration, centralizing integration of marine application modules, building a comprehensive and open marine application service management system, providing data sharing, information processing, decision support, and business operation for scientific research and industrial applications and other services.
3.2. Information acquisition layer

The information acquisition layer is the source of all the information of the digital ocean system, including static sources from various channels such as offshore surveying and mapping, island monitoring, underwater detection, marine fishery operations, marine buoy monitoring, marine scientific research, oil and gas platform environmental monitoring, remote sensing satellite monitoring, etc. Dynamic marine environmental information and marine business information [4]. This information comes from marine information systems of different departments and different business objectives. There are multiple types of video, audio, data, and documents. It is divided into structured, semi-structured and unstructured (mainly video and audio) data. With the continuous development of marine rights protection, marine resource development, marine scientific research, marine monitoring, marine management and control, and other businesses, these data will continue, rapid, and massive growth, which is the basis for the construction of a digital ocean system based on big data. as shown in picture 2.

Figure 1. Overall architecture of marine big data platform
3.3. Communication transport layer
The communication transmission layer is a way to transmit massive amounts of marine environment and business-related data to the ocean big data processing platform. It is also a shore-based communication station, ship-based platform, airborne platform, rights protection formation, space-based (satellite) platform, island monitoring, and oil and gas platform, Buoy platform and other information exchange channels between various marine business entities. Based on the characteristics of wide sea area and complex environment, the transmission methods are mainly wireless, including shortwave/ultrashort wave, communication satellites (maritime satellites, Zhong tong satellites), navigation satellites (GPS, Beidou), remote sensing satellites, laser communications, hydroacoustic communications (Underwater detection), 3G mobile communications, etc. Shore-based communications exist in a combination of wired and wireless networks.

3.4. Big Data Basic Platform of Digital Ocean
The big data basic platform of Digital Ocean is a basic computing platform (IAAS layer) based on cloud computing technology. Adopt virtualization technology to realize the virtualization of storage, computing, network and other facilities. Through cloud storage, virtualized network, virtual host services, and cloud platform management modules, it provides elastic, scalable, and on-demand big data centres for building digital oceans. Basic platform for service and high availability [5]. This kind of virtualization-based cloud computing basic platform technology and application are relatively mature. Typical products include VMWare system, Citrix's Xen-based virtualization system, and KVM-based virtualization system (see Figure 3).
3.5. Big Data Processing Platform of Digital Ocean

Digital Ocean's big data processing platform is a cloud computing platform (PAAS layer) based on a basic platform that adapts to the calculation and software development of massive data storage, analysis, and processing. Google, Amazon, IBM, Intel, and domestic companies such as Baidu, Alibaba, and Huawei all have corresponding commercial platforms. Intel and domestic big data processing platforms are mainly constructed based on Apache's Hadoop open source project, and Intel's Hadoop commercial platform is a representative of this type of platform. Hadoop is a software framework that can perform distributed processing of large amounts of data, and perform data processing in a reliable, efficient, and scalable way [6].

4. Risk assessment model based on the characteristics of marine disaster information

According to the Graham Law proposed by Graham and Kinney of the United States. The factors that affect the risk of environmental conditions are L (probability of a disaster), E (how often people are exposed to the dangerous environment) and C (the possible consequences once a disaster occurs). Use the product of the scores of these three factors to evaluate the risk of a single element of a disaster

\[ D_i = L_i + E_i + C_i \]  

In the formula: \( D_i \) is the single-element risk index of natural disasters, \( L_i \) is the average occurrence probability of each disaster factor; \( E_i \) is the time that the receptor is exposed to each disaster factor; \( C_i \) is the possible loss caused by each disaster factor, usually easy to use Destructive representation. According to the above-mentioned conditional attribute data, the rough set theory is used to perform attribute reduction and value reduction, eliminate redundant attributes, and find the decision factors closest to the current problem [7].

In a selected knowledge decision table \( S = (U, A, V, f) \), given an equivalence relation family \( P \subseteq A \), satisfying G is independent for any \( G \subseteq D \), and \( \text{IND}(G)=\text{IND}(P) \), then G is a reduction of P. Where \( \text{IND}(P) \) represents the indistinguishable relationship on \( U \), and the equivalence class of object x in \( U \) divided by attribute B is \( \text{INDB}(x) \).
\[ \text{IND}(P) = \left\{ (x, y) \in U^2 \mid \forall a \in P, a(x) = a(y) \right\} \]  

(2)

Define rough membership function

\[ \mu_x(x) = \frac{\text{IND}_x(x) \cap X}{\text{IND}_x(x)} \]  

(3)

By controlling the threshold of \( \mu_x(x) \), the scope of application of rough set can be expanded, so that incomplete knowledge in the data can be extracted and judged variable precision rules based on the calculation of upper and lower approximations and attribute reduction in rough set theory, and the rough set The membership function method is used to determine the credibility factor of the rule, thereby reducing the subjective factors of rule determination. In the decision table, \( U = \{ u_1, u_2, \ldots, u_n \} \) is the domain of discourse, and \( \omega \) represents the comprehensive effect information of disasters, disaster relief measures, and disaster losses that occurred in a certain area in the past. \( P(\omega \mid x) \) represents the probability that an object is in the state \( \omega \) under the description \([x]\). Through the above decision table model, this probability can be defined as the rough membership function \( P(\omega \mid x) = \frac{\omega \cap [x]}{[x]} \) of \([x]\), or the support of this decision rule.

5. Conclusion

The advent of the big data era does not depend on anyone’s will. Data-intensive science represented by big data will surely become the cornerstone of a new technological and industrial revolution. The construction and application of ocean big data platforms are just ocean informatization to adapt to big data. The inevitable development trend of the advent of the times. The IoT knowledge model for disaster prevention and mitigation system established in this paper divides and integrates disaster information data according to regional disaster theory and disaster chain theory, and extracts the rule knowledge implicit in this information according to the rough set method, so as to direct disaster prevention and mitigation Decision-making provides quantitative data support.

References

[1] Fang, S., Da Xu, L., Zhu, Y., Ahati, J., Pei, H., Yan, J., & Liu, Z. An integrated system for regional environmental monitoring and management based on internet of things. IEEE Transactions on Industrial Informatics, 10 (2) (2014) 1596 - 1605.

[2] Yang, C., Yu, M., Li, Y., Hu, F., Jiang, Y., Liu, Q., ... & Gu, J. Big Earth data analytics: A survey. Big Earth Data, 3 (2) (2019) 83 - 107.

[3] Fang, S., Xu, L., Zhu, Y., Liu, Y., Liu, Z., Pei, H., ... & Zhang, H. An integrated information system for snowmelt flood early-warning based on internet of things. Information Systems Frontiers, 17 (2) (2015) 321 - 335.

[4] Djedouboum, A. C., Abba Ari, A. A., Gueroui, A. M., Mohamadou, A., & Aliouat, Z. Big data collection in large-scale wireless sensor networks. Sensors, 18 (12) (2018) 4474.

[5] Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. Internet of things: Vision, applications and research challenges. Ad hoc networks, 10 (7) (2012) 1497 - 1516.

[6] Li, C. S., Darema, F., & Chang, V. Distributed behavior model orchestration in cognitive internet of things solution. Enterprise Information Systems, 12 (4) (2018) 414 - 434.

[7] Provost, F., & Fawcett, T. Data science and its relationship to big data and data-driven decision making. Big data, 1 (1) (2013) 51 - 59.