ARTICLE IN PRESS

Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Marine surveying and mapping system based on Cloud Computing and Internet of Things

Jiachen Yang [a], Chang Wang [a], Qiming Zhao [a], Bin Jiang [a,*], Zhihan Lv [b,*], Arun Kumar Sangaiah [c]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin, China
[b] School of Data Science and Software Engineering, Qingdao University, Qingdao, China
[c] School of Computing Science and Engineering, VIT University, Vellore, India

HIGHLIGHTS

- To deal with the complex and variable marine data, we propose a basic model of marine data from various sensors in order to make each sensor, equipment and system can share and process the data.
- To the best of our knowledge, there is no research focusing on the multi-sensors data in marine field. This paper propose a data processing algorithm of multi-sensors data in order to delete outliers and interpolate value.
- In previous studies, there is no algorithm or software for processing ship tracks systematically, the mathematical model we design will be beneficial to the route planning and marine surveying.

ARTICLE INFO

Article history:
Received 20 November 2017
Received in revised form 14 January 2018
Accepted 18 February 2018
Available online xxxx

Keywords:
Cloud Computing
Internet of Things
Marine surveying
Route planning

ABSTRACT

With the increasing popularity of Cloud Computing and Internet of Things(IoT), Cloud computing among with the IoT concept has become a hot research field in recent years. The ocean data information is more complex than the traditional data information. Based on the analysis of IoT in marine field, this paper will establish a simplified model of Cloud Computing on marine IoT and propose a ocean data processing algorithm. Specifically, different from traditional data processing algorithms, the proposed algorithm is based on the data of various sensors and cares deeply the effect of eliminating outliers and the smoothness of the processed curve. According to the processed data, we propose a mathematical model of route planning. In this model, the ship can arrive the destination line as soon and short distance as possible. Through analysis of experimental results by comparing the raw data and the processed data, it is shown that the proposed algorithm has a good optimization effect and the mathematical model is reasonable in actual measurement work.

1. Introduction

Cloud Computing is an increase, use, and delivery model for internet-related services. It takes computing to the pool of resources that are made up of a large number of computers, enabling a variety of applications to acquire the ability of calculation, storage space, and information services as needed [1,2]. Firstly, Cloud Computing allows users to use a variety of terminals at any location to obtain application services and the requested resource is from the cloud rather than the fixed entity. When measuring the marine data, the location and terminal are non-stationary, so the computation and storage of data will not be affected by the application of Cloud Computing [3,4]. Next, Cloud Computing has high scalability and large scale. The scale of the cloud can be extended dynamically and this dynamic extension is transparent to the user and does not affect the user's business and application [5]. Finally, it also has high reliability. Cloud Computing takes use of multiple copies of fault tolerance, multi-computing node isomorphism, and other measures to guarantee the high reliability of service [6,7].

With the purpose of improving the ability of system to collect data and analyze data, it is significant to combine the Cloud Computing with Internet of Things [8,9]. Internet of Things(IoT) is an important part of the new generation of information technology. It is still based on Internet and the expansion of Internet by thing to thing with the purpose of exchanging information and communicating [10,11]. There are various kinds of sensors(pressure,

* Corresponding authors.
E-mail addresses: jiangbin@tju.edu.cn (B. Jiang), lvzhihan@gmail.com (Z. Lv).

https://doi.org/10.1016/j.future.2018.02.032
0167-739X/© 2018 Elsevier B.V. All rights reserved.

temperature, speed, etc.) in IoT, allowing us to use the data of each sensors and make the sensors interact each other [12,13].

As the basic model of Cloud Computing and IoT shown in Fig. 1, there are three parts in this model. The various sensors collect data and interact with each other. Then the data will be transmitted to the cloud. Based on Internet server and cloud, the Cloud Computing is used to deal with the larger quantity of data. According to the processed data, they will be used in many applications [14,15].

According to IoT and Cloud Computing in marine field, it is a new technology in IoT, which can be used to process the big data and develop new oceanic technologies [16,17]. It combines the IoT and ocean. To be more specific, IoT can be linked with a large number of information from various sensors [18]. According to the contract agreement, any items connected with the Internet in order to achieve intelligent identification, positioning, tracking, monitoring and management of a network [19]. The basic feature of IoT is the comprehensive perception, reliable transmission and intelligent processing of information, the core of which is the information interaction between people and things or different devices [20]. Wherefore, in view of the application of large-scale wireless sensor network, it is necessary to study the various problems involved in IoT and Cloud Computing and verify the conclusions of theoretical research through the actual data processing system and the marine surveying and mapping system [21,22].

At present, oceanic information is usually acquired by using measuring vessels, satellites and oceanographic stations. However, with the development of the oceans and the increasing demand for data, IoT-based data acquisition and cloud-based data processing can improve data collection efficiency, solve the problem of low data utilization and also can be applied to the marine mapping system. But the problems encountered mainly in the following two aspects, one is in data processing that the application of big data processing in other fields is mature, but relatively few in the field of the ocean, which needs a suitable data processing algorithm; the other is that in the actual surveying and mapping of marine survey ships, it sometimes needs to follow the designated route, so a basic model of route planning is necessary.

Marine charting is an accurate measurement and description of all kinds of data information of the ocean, such as water flow velocity, wind direction, wind speed and other data information. And these data information is analyzed and processed in order to set up the voyage line [23]. For the ship underway measurement, the basic layout is in accordance with the planned survey to make an underway measurement. The measuring ships need to change their direction frequently and the methods let these ships sail on many different lines [24]. Thus, the establishment of a navigation system for sea ships is based on the analysis of the minimum turning radius and the influence of wind direction, water flow velocity and the direction of the ship in order to reach the route in the fastest and shortest way. Based on the data information and the relative position of the ship and the measured line, we establish the mathematical model of the route planning of the navigation to the test line under different conditions by means of analysis and simulation [25,26].

In this paper, we propose a basic model of Cloud Computing in marine IoT. And according to the model, we put forward the data processing algorithm based on the data interaction between various kinds of sensors and propose a mathematical model of route planning according to the processed data. According to the Could Computing and marine IoT, there are three contributions in this paper. First of all, this paper proposes a basic model of Cloud Computing and marine data from various sensors in order to make each sensor, equipment and system can share and process the data. In addition, a data processing algorithm of multi-sensors data is proposed in order to delete outliers and interpolate value. At last, the mathematical model we design will be beneficial to the route planning and marine surveying.

The rest of the paper is organized as follows: Section 2 is devoted to the introduction of the related works. Section 3 illustrates the proposed approach. Experimental results as well as the evaluation of the proposed approach are presented in Section 4. Ultimately, Section 5 concludes the paper.

2. Background and motivation

Most marine detection systems are made up of payloads, detection equipment, launchers and ground stations. And the station mainly includes the ground receiving equipment that receives the measurement information, tracking and detecting equipment, positioning speed measuring equipment and data processing system. The data processing system of each monitoring equipment stores, processes and visualizes the collected data. Ground handling systems in early stage have the drawbacks of low storage capacity, low computational speed, etc. With the development of computer technology, real-time information and communication efficiency has been greatly improved by applying various data processing software and a stable and reliable database [27].

Data processing plays a key role in Marine Mapping System. The data acquisition can automatically collect the non-electric or electric signal in the unit, such as the sensor and other test equipment, etc. [28,29], and send them to the upper machine for analyzing and processing. Data processing module will standardize and visualize data to provide data support for the system [30,31].

Compared with the existing data processing methods, the Marine Surveying and Mapping System based on Cloud Computing and IoT combines the characteristics of the data acquisition of the detection equipment. Based on the analysis of the complicated marine environment, the reasons for the measurement error of the sounding rocket are summarized. What is more, the system errors are corrected to make them more realistic to reflect the real situation of the marine atmosphere. And it also breaks the traditional telemetry data processing mode and solves the problem that most marine wireless sensors or equipment have poor processing flexibility and low utilization rate for collecting environment parameters such as temperature, humidity and wind direction.

2.1. Data preprocessing in data acquisition module

The data acquisition is a kind of indirect access technology, such as the marine environment parameters, containing temperature, air pressure and other marine parameters, which is transformed into electrical signals through a certain transformation technology [32]. After data acquisition, the data information will be transmitted to the cloud. With a large quantity of data, it is necessary to process data in the cloud. In multiple conversions, the data will be mixed with a small amount of noise data, which affect the accuracy of the data. In order to reduce or prevent the influence of noise, digital filtering technology can be used to eliminate noise and improve the accuracy of data [33].

In the data processing methods, data filtering is a crucial step, while most of exiting methods are using mean filtering method. The mean filtering is the average of the sum of the data in the sampling sequence $\{X_i\}$, the formula is as follows:

$$X = \frac{1}{n+1}\sum_{i=0}^{n} X_i \tag{1}$$

The mean filtering is very effective for filtering the random interference signals mixed in the detected signals. The signal is characterized by an average signal that moves up and down near a numerical range. But the filtering result of this method can be affected when the noise data is uncertain [34].

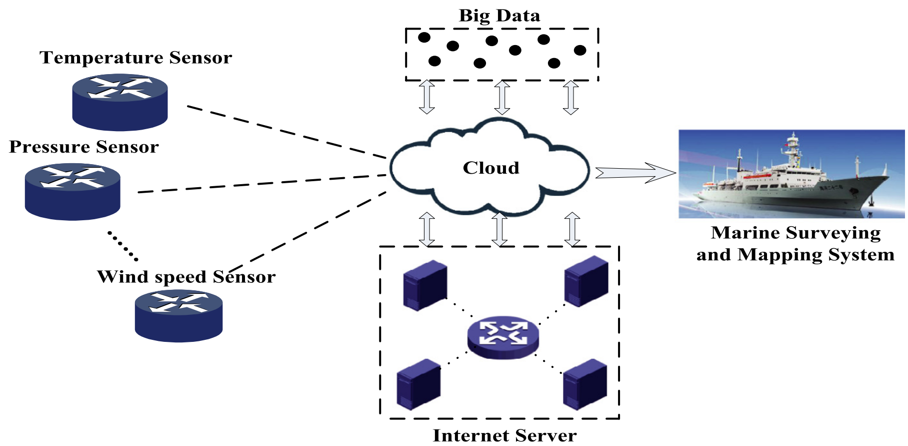


Fig. 1. Basic model of Cloud Computing on the Internet of Things.

The median filter method is used to treat the abnormal value of data [35]. Suppose the data sequence as $X_0(i)$ where $i = 1, 2, \ldots, N$.

Take the element j as the starting point, and take out $(m+1)$ elements in order. In the formula, m is even and j takes the value $1, 2, \ldots, N-m$, the data of each participating median filter is $(m+1)$, so the sequence is rearranged into a new sequence in order of magnitude $X_0(j), \ldots, X_0(j+m)$. Each j takes the middle elements in the sequence to form a new sequence and this new sequence is the result of $(m+1)$-point median filtering: $X_1\left(j+\frac{m}{2}\right)$ where $j = 1, 2, \ldots, N-m$

The median filter method must be given a test value that is not physically intuitive, so we will optimize the method for dealing with the abnormal data in part three [36,37].

2.2. Route planning methods in marine field

The measurement of the ocean is very different from land measurement, it is that the survey line of measurement will be influenced by many factors such as wind and wave. Even though the operators try to keep the ship on a preset plan, the actual flight path is a curve that moves around the planned course. Let us take the water depth measurement as an example [38]. At present, the expression of water depth data is in the form of water depth, where the location of the water depth can be represented in either direct location point or interpolated points. In order to compensate the system error of the test line network, we can use the two points connected on the main test line to find the intersection point of the nearest two points on the line, and then make the system error compensation. This method will be more effective in the situation of small scale or straight survey line. In the measurement of large scale depth, it is likely to bring the measured line noise into the intersection when the intersection point method is obtained by using the straight line. Next, this paper summarizes the mathematical model based on the observation position and the calculation of the ship position, the random disturbance of navigation and the smooth trajectory, and the adjustment calculation is carried out [39].

In the ship's route planning, most of the technologies are now focused on the algorithms in positioning and route correction. But there are few algorithms for route layout, it is necessary to build a suitable model of route planning in order to solve some problems in practical measurement. Also, the combination of the route planning model and route correction algorithms is very important [40].

There are two ways to measure the position of the ship. The first is a method of measurement, which makes use of the GPS positioning system, optical or acoustic system, etc. and observes several given points to determine the position coordinates on each measurement parameter point [41].

When using multiple sensor measurements, k represents the number of locating sensor, so the observation i measured by the sensor k on the measuring point j is $U(i{\cdot}j)k$. Also the power is $U(i{\cdot}j)k$ and $P(i \cdot j) \cdot k$. The corresponding error equation is:

$$V(i \cdot j)k = (a_{i \cdot j}\delta x_j + b_{i \cdot j} \cdot \delta y_j - l_{i \cdot j}) \cdot k, P(i \cdot j) \cdot k \tag{2}$$

The error equation of the whole measuring line can be written:

$$V' = A \cdot X - l, P_A \tag{3}$$

Next we can calculate the error equation of the reckoning method. The position of the $(j+1)$ is obtained by the j coordinate point and the heading and speed of that position. Take the result to the error equation of the measurement method, and the constant term is equal to zero. Then the error equation can be calculated as follows:

$$\left.\begin{aligned} V'_{x_{i \cdot j+1}} &= \delta x_{j+1} - \delta x_j + \eta_x \\ V_{y_{i \cdot j+1}} &= \delta y_{j+1} - \delta y_j \end{aligned}\right\} \tag{4}$$

Assuming that $V' - \eta = V$, then we can convert expression (4) into the following equation:

$$\left.\begin{aligned} V_{x_{i \cdot j+1}} &= \delta x_{j+1} - \delta x_j \\ V_{y_{i \cdot j+1}} &= \delta y_{j+1} - \delta y_j \end{aligned}\right\} \tag{5}$$

The calculation error equation of the whole line is written:

$$V'' = B \cdot X, P_B \tag{6}$$

By combining the two error formulas of the measurement method and rocking method, the basic equation of the adjusting of survey line can be represented as (7). When the ship is affected by a variety of dynamic factors, such as the wind and wave, it will generate random errors to the heading and speed as a result of forming the error source.

$$(A^T P_A A + B^T P_B B) \cdot X = A^T P_A l \tag{7}$$

In Fig. 2, x_i represents the measurement of the fixed position, and y_i means that the position of the estimated method z_i is the position of the adjustment. There will be some observation errors when we fix the position by measurement, and the calculated position contains the random error causing by the heading or the speed of the ship when we use the reckoning method. Therefore, the adjustment point calculated by the methods is the best position z_i.

Please cite this article in press as: J. Yang, et al., Marine surveying and mapping system based on Cloud Computing and Internet of Things, Future Generation Computer Systems (2018), https://doi.org/10.1016/j.future.2018.02.032.

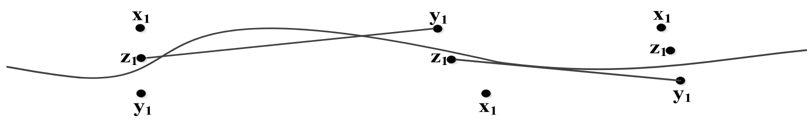


Fig. 2. Illustration of reckoning method.

3. Proposed method

According to the information data from various kinds of sensors, it can be known that the information transmission and data processing between sensors is significant.

3.1. Basic model of data transmission in marine Internet of Things

In the oceanic data system, there are many devices and sensors. They previously interact with the data and eventually converge to the receiver. In this paper, the main discussion is about atmospheric data collection, control and processing in the cloud. And we also analyze the interaction between various devices and sensors and propose a model of data transmission. The model of information transmission in marine IoT is shown in Fig. 3.

In Fig. 3, the data information from one sensor interact the other sensors. The various sensors collect data information and transmit the data to transducer. Then the transducer can change the data into signal in order to make the receiving terminal receive data information. The receiving data information will be transmitted to the cloud. In the cloud, data processing method will be used to deal with data.

3.2. Sounding rocket data processing

Low-altitude sounding rocket is a device used for detecting marine environment parameter of the atmospheric boundary layer and mixed layer, and also utilizes the sensor carried by the radiosonde to detect air temperature, air pressure, relative humidity and other atmospheric environmental parameters. The low-altitude sounding rocket is usually launched on a ship, with a rocket launch altitude of about 1000 m high. After reaching the highest point, the parachute opens and the rocket sends data stably once per second. It has the characteristics of high spatial resolution, affected lightly by the hull environment and it can be used in all-weather situations. The working model of rocket is shown on Fig. 4.

The performance index of each sensor is shown in Table 1. Due to the movement of the sounding rocket, the launch of the sonde and the failing process strongly influenced by sea conditions and atmospheric environment, the data usually show abnormal condition. After a large number of sounding rocket data processing and analysis, we summarized the abnormal situation shown in Table 2.

Through the analysis of a large number of measured and abnormal data, we sum up the following quality control methods. After transmitting data to the cloud, a suitable data processing is significant. In the data processing stage, there are three kinds of sensor data for each observation layer. A sensor data may be abnormal in a probe layer, while the others are normal. Then if we delete the data of this layer, it will cause the normal data to be deleted. In this situation, we mark the bad data rather than delete it.

(1) The extraction of the falling stage.

In the initial stage, the launch of the sounding rocket sensor will produce detection error because of failing to adapt to the environment and this error usually occurs near the highest point. Only when the rockets through the highest point and the parachute opened, the data is valid. Therefore, it is necessary to judge the point of the falling stage, which is carried out on the basis of the data scope qualification.

Normally, the pressure change linearly with height in the vertical direction and the highest point of the rocket should be the lowest point of atmospheric pressure. So the pressure gradient testing method can be used to judge data in the decline stage.

Taking the sum of five consecutive data points A_1 , A_2, A_3 into account, shown as Fig. 5, if the data point is empty, then we take the point behind the empty data. The following is the discussion of the gradient between A_1 , A_2, A_3.

When $A_1 > A_2$, if $A_2 < A_3$, then A_2 is the lowest pressure and the midpoint of A_2 is the starting point of the descent stage; if $A_2 > A_3$, then the pressure is decreasing, so the data for the rise phase continues to be looked down. And when $A_1 < A_2$, if $A_2 < A_3$, then the pressure is increasing and the value of A_1 is the starting point of the descending phase; if $A_2 > A_3$, then there is non-increasing or non-increment, there are data errors in A_2.

If the highest point is not found, the starting point of the value move one data points backwards and loop operation until the highest point is found.

(2) Determine the stable point.

When the rocket reaches its highest point, the parachute will deploy and the rockets radiosonde will start to fall at a constant speed. At this time, the sensor detection data is stable and effective, so after finding the descent phase point we should continue to find stable point. The selection of the stable point is carried out in two steps.

Find the empirical stability point. First of all, according to the accuracy of the pressure sensor, we assume that the mean square deviation of pressure of the adjacent points is 1 hPa and the difference between two consecutive points is calculated from the highest point. When the difference between the 10 successive points is less than 1hPa, the starting point of the continuous point will be considered as the stable point of the instrument. If the stable point is not found, though moving down the starting point and searching for loops, we will set the point as the empirical stability point. The data sequence is supposed as $X_0\,(i)$ where $i = 1, 2, \ldots, N$. And when $i = 1$ we call the point is the empirical stability point.

Recount the mean square deviation of the data sequence. Let k equals to the m times of mean square deviation, the empirical stability point is regarded as the starting point and we compare the difference between adjacent points and k. If the value of 10 consecutive points is less than k, then the $X_0\,(i)$ point is the stable starting point. If there is a point greater than k, then the starting point is moved one point backward and so on.

If the whole data sequence does not find the starting point, then we increase the k value while reducing the number of continuous points, in accordance with the above method to do the search.

3.3. Data processing in marine internet of things

As for a large quantity of data in oceanic data system, there will inevitably be a lot of error data in the system. Therefore, this paper analyzes and processes the data, and proposes a basic data processing algorithm.

The steps of the algorithm are shown in Fig. 6.

(1) Initial check.

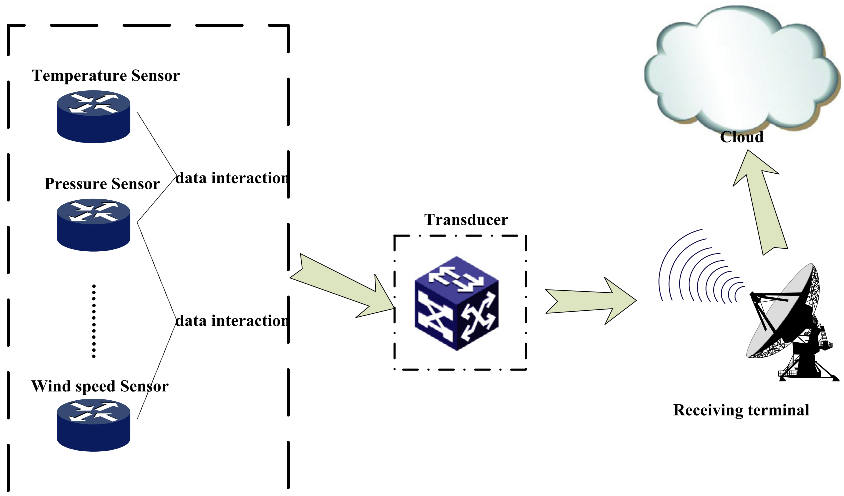


Fig. 3. Model of data interaction in marine Internet of Things.

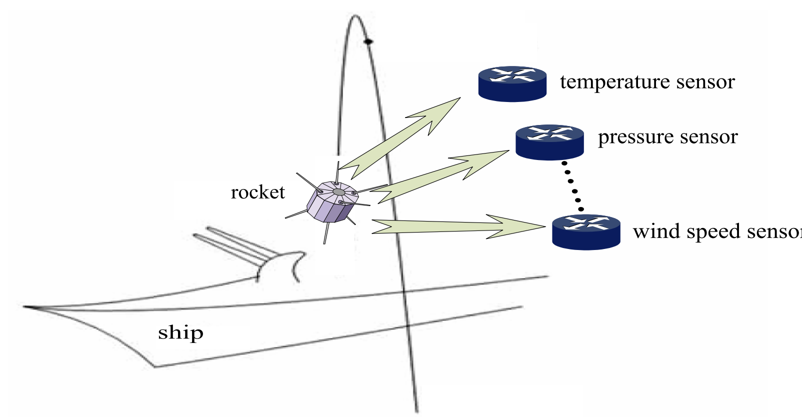


Fig. 4. Working model of sounding rocket.

Table 1
The performance index of each sensor.

Sensor	Type	Precision	Resolution	Induction time	Detection range
Pressure	capacitor	1.0 hPa	0.01 hPa	<0.1 s	600–1050 hPa
Temperature	Bead thermistor	0.3	0.01	>0.1 s	−55–50
Humidity	Humidity capacitance polymer	3%	0.1%	<1.0 s	0–100%

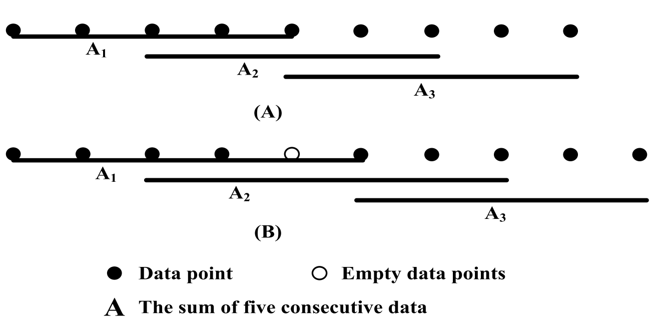


Fig. 5. Schematic diagram of smoothing data.

Since the ship is a navigational observation, it is necessary to get the common signal in order to record the observation position and correct element value. In the preprocessing stage, the devices receive the observed data from the IoT in real time and perform operations such as unpacking, error handling and format conversion, and send the preprocessed data to the subsequent units for further processing. Besides, when the common signal is input abnormally, the GPS information and the compass information are lost, so that the current position information cannot be recorded. Taking the signal detection from the signal access terminal and the current common signal comparison into account, we delete this data if the detection is not normal input.

(2) Mark outliers.

As for the outliers that appear in the data set, this method takes the way of marking. An outlier refers to an extremely unreasonable data value that deviates from most other data either out of the range of a device's measurement or a non-real mutation. There are more than ten sensors at each observation point of the observation station to record the data. For an outlier from one sensor, if take

Table 2
The data exception analysis table of the sounding rocket.

Order	Data characteristics	Possible reasons
1	Incomplete data	Sensor fault
2	Too much data, starting point is not barometric minimum	Premature launch of data transmitter
3	A lot of barometric data is obviously wrong (for example, P >2000 hPa)	Pressure sensor fault
4	The pressure gradient is unstable	Rocket fault
5	The data is incomplete a nd the minimum barometric pressure is greater than 1000 hPa	Parachute opening failed
6	The data is normal but less, and the minimum barometric pressure is too large	The launch of the rocket is not vertical
7	Incomplete data, or prompt at the end of the file	Rocket fault
8	The air pressure is not increasing sequentially and the data acquisition is uneven	Influence of low altitude disturbance airflow
9	Humidity value greater than 100%	Sensors into the cloud
10	The quality of the is not stable	Sea conditions or signal interference signal interference

the way of deleting data, it is bound to affect the other elements of this point and make all data of this point to be regarded as outliers. It is necessary to mark the outliers to avoid spreading to adjacent sensors, so that the original reasonable observations are not affected.

The specific approach is setting a reasonable fluctuation value of observation data elements according to the specific circumstances of the observation point. So comparing the data with the upper and lower limits of the reasonable value, the data over the range of reasonable value as an outlier should be marked. After marking outliers, there are n data value of an element as $X_i\,(i = 1, 2, 3 \cdots n)$. And for each element of observation point, the average value is $\overline{X}$. By using a moving average method, the standard deviation can be expressed as:

$$Y = \sqrt{\sum_{i=1}^{N} (X_i - \overline{X})^2/(N-1)} \tag{8}$$

According to the set value m, X_i will be considered as an outlier to be marked if $\left|X_i - \overline{X}\right| > mY$.

(3) Consistency check.

In the marine system, the distribution of the elements is divided into vertical and horizontal directions and it is generally continuous in the vertical direction. Therefore, it is necessary to detect whether the observed value of different isosurface satisfy the consistency of the change. Assume that the change of a element in the vertical direction is linear, the difference between two layers can be expressed as:

$$DX = \left|X_i - X_j\right| \tag{9}$$

Take the wind speed for examples, the difference between two layers (DX) should set a reasonable value. In this paper, the DX value is set to two. Then we choose speed value of three layers on behalf of the bottom, middle and top data, represented as X_{1i}, $X_{2i}, X_{3i}\ \ (i = 1, \ldots, M)$. Through the expression (8), the standard deviation of the bottom layer and the middle layer is calculated as the following expression:

$$Y_1 = \sqrt{\sum_{i=1}^{M} (X_{1i} - X_{2i})^2/(M-1)} \tag{10}$$

As the same way, the standard deviation of the bottom layer and the top layer, the bottom layer and the top layer can be calculated as Y_2, Y_3. Then we check the consistency according to the standard

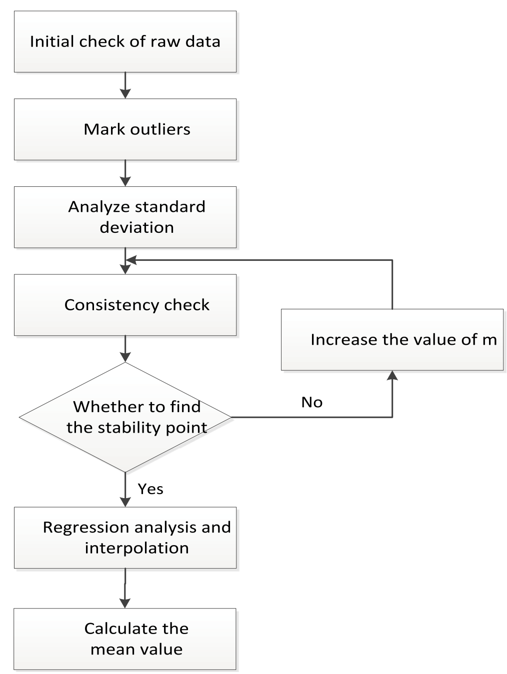


Fig. 6. Frameworks of the data processing method.

deviation. Comparing the standard deviation and the difference of each point, only when the value of difference is smaller than the value of standard deviation then the data meets the requirements of consistency. That is, the data meets the following inequality group:

$$\begin{cases} |X_{1i} - X_{2i}| < Y_1 \\ |X_{2i} - X_{3i}| < Y_2 \\ |X_{1i} - X_{3i}| < Y_3 \end{cases} \tag{11}$$

(4) Regression analysis and interpolation.

Through the analysis of a large number of observation data, it is found that the detection data from sensor is seriously affected by the environment of hull, such as a layer of temperature, wind speed data in a certain period of time significantly deviated from the other two observations. In this situation, the regression interpolation method is used to deal with the marked outliers. Supposed that the number of experiment data group is m and

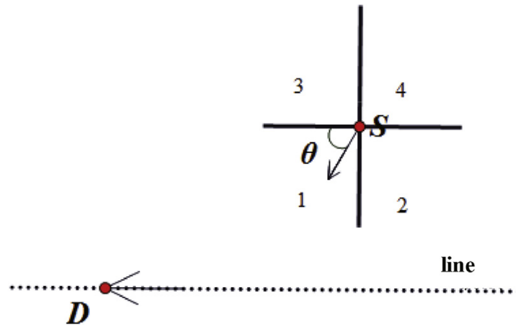


Fig. 7. Classification diagram according to the angle.

the relationship between each group is linear. According to the least squares method, the expression of regression function and measured data can be defined as:

$$Q(b_{0,}b_1, b_2) = \sum_{i=1}^{m}(y_i - (b_0 + b_1x_1 + b_2x_2))^2 \tag{12}$$

To obtain the minimum value of the above expression (12), the coefficient should satisfy the following equations:

$$\frac{\partial Q}{\partial b_0} = 0, \frac{\partial Q}{\partial b_1} = 0, \frac{\partial Q}{\partial b_2} = 0 \tag{13}$$

According the method of solving the regression problem, the solution to this equation can be calculated. And it is able to get the regression function by replacing the value of each coefficient. The regression function is used to fit the data and interpolation.

(5) Regression analysis and interpolation.

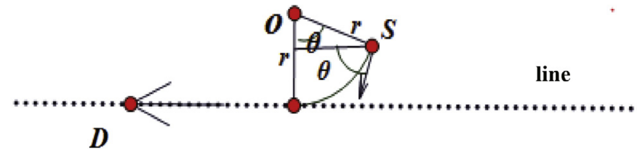
Considering the unit time of data collection, take the temperature sensor for example, the unit time is three seconds. But in actual measurement, the average data per minutes will be more convenient and effective. According to the data after processing, the average value of the data except outliers can be calculated.

3.4. Mathematical model of route planning

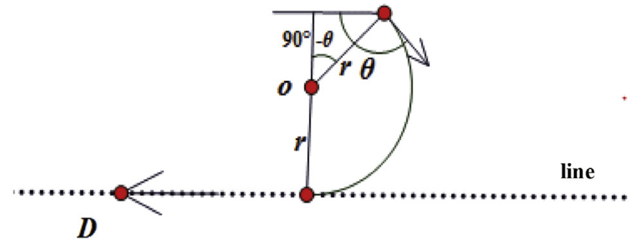
According to the data transmission and data processing in the cloud, it is known that the device get the reasonable data. As for the data of wind speed and flow velocity, they can be widely used in various aspects. And the complexity of seabed terrain can be determined by the observation data. Considering the influence of wind speed and flow velocity, the route of ship will change compared to the ideal situation. And the complexity of seabed terrain is also the important factor of route. In this situation, a mathematical model of route planning based on GPS is proposed in this paper with the purpose of making the ship on the established route.

The ocean survey line consists of the main line and the inspection line. The definition of tracing planning is searching a route from the main line to the inspection line. But in practical applications, for the survey ship, the current problem is finding a route from the current position to a straight line with the direction. Depending on the particular situation, there will be many ways of route layout. And according to different situations, the way of route layout is more various and complex. In this paper, the mathematical model is divided into three situations.

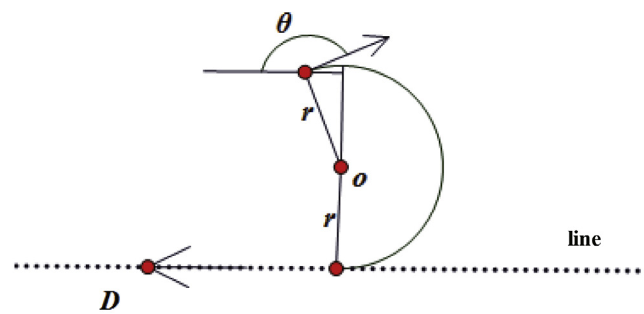
In this model, the following parameters are defined. The starting position of the ship is S and the position of arriving the line is D. Also the distance between point S and point D is W. In other word, this model makes a route from S to D with shorter distance. According to the performance of the ship, the minimum turning



(a) Illustration of $0 < \theta < 90$.



(b) Illustration of $90 < \theta < 180$.



(c) Illustration of $180 < \theta < 270$.

Fig. 8. Illustration of basic model.

radius of the ship is defined as R_0. What is more, the angle between the direction of the ship and the direction of the line is θ.

(1) The Basic situation.

The basic situation is when the distance between the point S and the point D is small. Under this circumstance, the ship can go ahead on the line, because it will provide more time and opportunity to adjust the direction of ship while the ship is in the actual measurement operation. But considering the situation that distance between the ship and the line is smaller than R_0, the ship reach the line before the point D. But from a modeling point of view, in order to get a more general case, this critical situation allowed to consider later. So make the following assumptions:

a. The ship arrive the line as soon and short as possible.

b. The distance between the ship and the line is smaller than $2R_0$.

c. The distance W is big enough.

Under these assumptions, the basic situation can be divided into four parts in consideration of θ, which is shown in Fig. 7. According to the θ, the ship can directly reach the line when the θ in range of 0 to 270. The route of each situation is show in Fig. 8.

Though comprehensive analysis, the track radius in above conditions can be expressed as:

$$R = h/(1 - \cos\theta), R > R_0 \tag{14}$$

where h represents the distance between the ship and the line. When the R satisfies the expression (14), the ship can find a suitable route in Fig. 8. However, it is unacceptable that R is too long. So a limitation factor and threshold should be added. In this model, the value of threshold is set to $4R_0$. Take the θ between 0 and 180 for example, the expression (14) can be changed as a inequality group

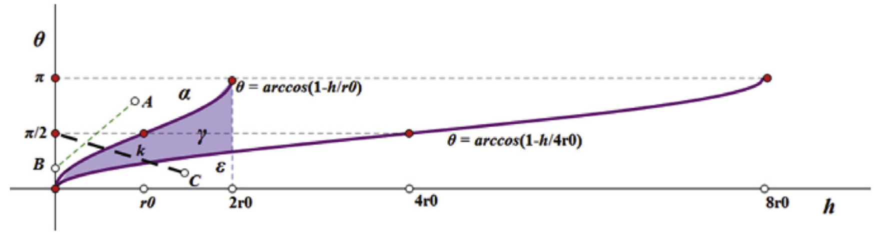


Fig. 9. Illustration of feasible domain.

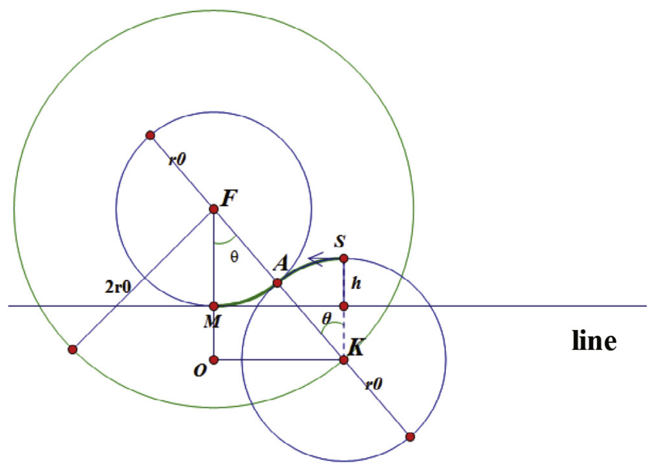


Fig. 10. Illustration of the situation that angle is 0.

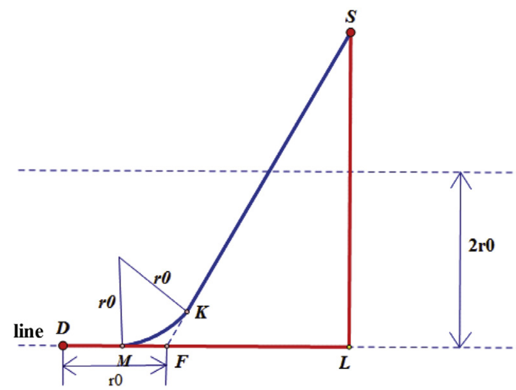


Fig. 11. Illustration of the situation that W is too long.

based on the threshold. It is expressed as:

$$\begin{cases} \theta \leq \arccos(1 - h/R_0) \\ \theta \geq \arccos(1 - h/4R_0) \end{cases} \tag{15}$$

The relationship between h and θ is shown in Fig. 9 According to Fig. 9, the feasible domain is filled with purple, represented as γ. When θ is in domain ε, it is necessary to decrease the value of angle. But the angle is limited by the minimum turning radius R_0, the relationship between h and θ is continuous. The straight line AB indicates the change qualitatively according to Fig. 9. It is obvious that the line AB will not go through the feasible domain γ. When θ is in domain ε, that is the R is bigger than $4R_0$, the method of rotation in order to increase the θ is reasonable. As the dotted line from point C in Fig. 9, the route can go through the feasible domain γ.

Considering an extreme case where $\theta = 0$, $h > 0$, the radius R is infinite according to expression (14). It is necessary to find a new route for this situation. In this model, the ship will have an anticlockwise rotation firstly as a result of changing the situation to the basic situation that θ between 0 and 90. It is be shown in Fig. 10.

This situation is also suitable for the situation that θ between 270 and 360. The ship only needs to have a clockwise rotation to make θ equal zero at first. And then the following steps are same to the situation that θ is zero.

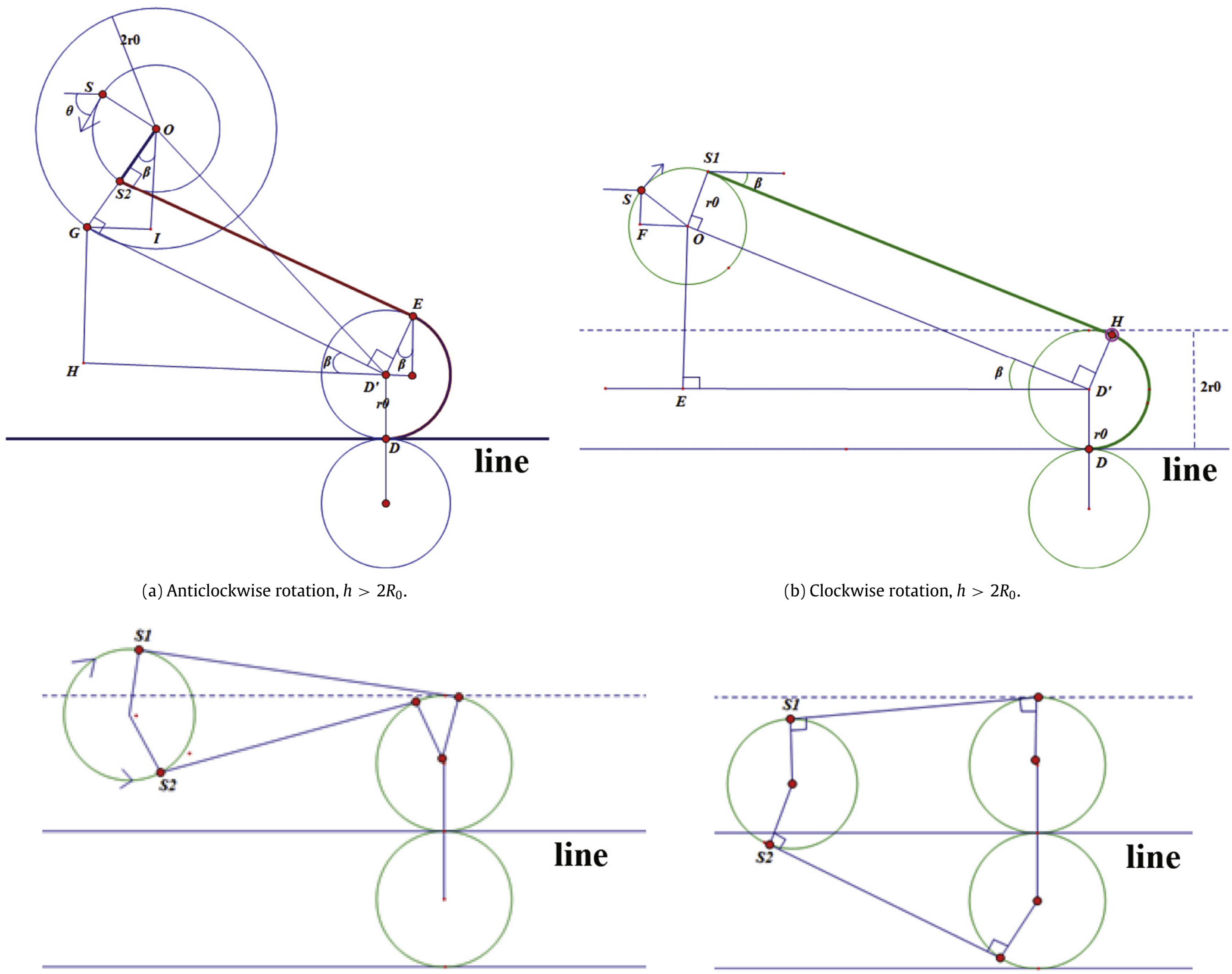
(2) The situation that W is too long.

In this model, take the situation that the W is much bigger than R_0 and $h > 2R_0$ for example, the other situations are based on this situation. Comparing the route of arriving the line at first, represented as route SLMD in Fig. 11 and the route of going straight in the direction to get the nearer point, represented as rout SKMD in Fig. 11, it is obvious that the distance of SKMD is smaller than the distance of SLMD. The way of route SKMD is to make the ship close to the point D, so that the situation is converted to the basic situation that is between 0 and 90. It should be noted that the horizontal distance between the ship and the line-up point needs to be taken into account in this case. And the horizontal distance is gradually reduced in the process from the far point to the near point. It is necessary to arrange that the ship starts going on line before point D.

(3) The situation that the ship in the upper part of the line.

Based on the situation of the (1) and (2), this model can be converted to a simple model that divided into two situations according to the θ. When the θ is between 0 and 180, the ship can make a anticlockwise rotation at first and then go straight to get nearer to the point D, which is shown in Fig. 12(a). As we can see from the figure, the ship get nearer point E and the situation can be converted to the basic model. And the situation when the θ is between 0 and 180 is similar to the above situation. As shown in Fig. 10(b), the ship make a clockwise rotation and then get the nearer point E. Above two situations are built in the case of $h > 2R_0$, it is necessary to discuss the case of $h < 2R_0$. When the distance between the ship and line is too smaller, the ship cannot turn around to the line directly. In this model, a route designed according to the h is shown in Fig. 12. The Fig. 10(c) represent the situation that the ship is in the upper field of the line but $h < 2R_0$. Similarly, Fig. 10(d) represent the situation that the ship is in the bottom field of the line but very close to the line.

In summary, the mathematical model of route planning is proposed to solve the most situations in the route planning problem. After analysis and verification, the model is suitable and reasonable. But the model does not take into account all the critical situation, resulting in appearing some problem in the actual travel of the ship. And take the factor that influence the route of ship into account, such as wind speed and flow velocity, it is necessary to build a correction method to reduce the impact of these factors. The value of these factors can be collected from various kinds of sensors. The ship can get the data value finally through data transmission and data processing. According to the value of wind



(a) Anticlockwise rotation, $h > 2R_0$.

(b) Clockwise rotation, $h > 2R_0$.

(c) $R_0 < h < 2R_0$.

(d) S is near to the line, $h < 2R_0$.

Fig. 12. Illustration of the situation that ship is above on the line.

speed and flow velocity, the ship can reduce the impact through existing correction algorithms.

4. Experimental result

In this paper, the data such as pressure, temperature and wind speed received directly by the ground receiving equipment without any manual or computer automatic quality control are stored in the original database. The raw data is observed for real-time storage. At the same time, sounding and wind-measuring data within the first 5 min are stored and the data storage format and content are standardized according to the measurement requirements. When the measurement is finished, the basic data files are converted from observed raw data into meteorological elements and stored together with the basic parameters of the station, the merit determination and instantaneous ground meteorological element values. After the quality control, the methods remove the obvious error value, fit the curve and smooth the curve according to the normal trend of the data curve. The required parameters and data storage format in the above process are in accordance with the unified meteorological standards.

The real-time processing of ocean data is implemented on minicomputers that are equipped with radars. Because minicomputers have small memory capacity and low computational speed, they cannot process sounding data with high code rate and data cannot be saved in time. Then the post-processing requires manual typing, which makes it difficult for subsequent analysis and research. To overcome the drawbacks of early processing systems, by using the data collection and processing of distributed dual processing approach (that is, the use of single-chip data acquisition, AT machine to do real-time processing), we can get the complete measurement data and curve. In recent years, the processing of oceanic data has achieved rapid development. Although various processing systems have perfect functions, their versatility is not broad and the reason is that different projects have different requirements for loading, processing data and formatting. In order to solve these problems, based on the models and the algorithms section3 proposed, we proposed a Marine Surveying and Mapping System via Cloud and IoT. In this model, the scale of data collection is improved through the interaction and data transmission of multi-sensor based on IoT, and the coupling between different types of data is also enhanced. What is more, the data processing algorithm based on cloud computing improves the storage capacity and the operation speed. And we also proposed a model of route planning according to the processed data. The comparison of different data processing methods is shown in Table 3.

Table 3
Comparison of data processing methods.

	Minicomputer processing	Distributed dual processing	Computer real-time processing	Processing via Cloud and IoT
Memory capacity	Small	Middle	Middle	Large
Calculating speed	Low	Middle	Middle	High
Reliability	General	General	More	Most
Security	Small	Middle	High	High
Sharing	Low	Middle	Middle	High
Maintainability	General	General	More	Most

Table 4
Raw data of wind speed.

Order	Direction (deg)	Speed (m/s)	Altitude (m)
1	329.0	2.80	35.0
2	330.3	2.93	39.5
…	…	…	…
29	341.2	4.85	149.5
30	340.8	4.84	153.0
31	340.5	4.74	157.6
32	340.3	4.84	156.3
33	339.9	4.84	159.4
…	…	…	…
90	7.6	5.02	394.5
91	7.9	5.12	398.7
92	8.1	5.76	401.0
93	8.3	5.23	402.9
93	8.8	5.35	407.0
…	…	…	…
104	15.9	6.34	445.8
105	16.3	6.42	450.0
…	…	…	…

Table 5
Processed data of wind speed.

Order	Direction (deg)	Speed (m/s)	Altitude (m)
1	329.0	2.80	35.0
2	330.3	2.93	39.5
…	…	…	…
29	341.2	4.85	149.5
30	340.8	4.84	153.0
31	340.3	4.84	156.3
32	339.9	4.84	159.4
…	…	…	…
89	7.6	5.02	394.5
90	7.9	5.12	398.7
91	8.3	5.23	402.9
92	8.8	5.35	407.0
…	…	…	…
102	15.9	6.34	445.8
103	16.3	6.42	450.0
…	…	…	…

Table 6
Raw data of temperature.

Order	Temperature (°C)	Altitude (m)
1	11.47	35.0
2	11.11	39.5
…	…	…
65	9.41	293.2
66	9.40	297.1
67	9.39	300.4
68	9.37	301.1
69	9.35	305.1
…	…	…
89	9.05	394.5
90	9.04	398.7
91	8.92	401.1
92	9.02	402.9
…	…	…
105	8.62	445.8
106	8.60	450.0
…	…	…

Table 7
Processed data of temperature.

Order	Temperature (°C)	Altitude (m)
1	11.47	35.0
2	11.11	39.5
…	…	…
65	9.41	293.2
66	9.40	297.1
67	9.37	301.1
68	9.35	305.1
…	…	…
88	9.05	394.5
89	9.04	398.7
90	9.02	402.9
…	…	…
103	8.62	445.8
104	8.60	450.0
…	…	…

Based on the models and the algorithms section3 proposed, marine surveying and mapping system contains three major steps. Data interaction of various sensors and data transmission is the first step. Then data processing of the raw data is important. The system will plan route according to the processed data. In the experiment, in order to prove that the proposed data processing algorithm has a good optimization effect, we take the data from actual measurement into account. Through the experimental simulation to evaluate the feasibility of the proposed methods, take the particularity of the experiment into account, we choose the parameter such as smoothness, reliability, accuracy and adaptability to evaluate the proposed methods. Take two groups data for example, the first group data of wind speed is shown in Table 4.

According to the proposed algorithm, we use computer for programming to verify the feasibility of the algorithm and draw the image according to the raw data and the processed data. And the processed data is shown in Table 5. The Fig. 13 shows the graph of the data before processing and after progressing.

In Fig. 13, the difference between the curve of raw data and the curve of processed data is obvious. As for the curve of raw data, it is obvious that the curve have catastrophe points. After data processing based on proposed algorithm, the curve is smoother and the catastrophe points are removed. Also through the method of curve fitting and interpolation in this algorithm, the new interpolation point is added into the curve according to the density of the data. While the data distribution is dense, the algorithm will delete the outlier value rather than add an interpolation point into the curve.

In order to verify the algorithm more convincingly, we choose the second group of temperature data from temperature sensor. The group of raw data is shown in Table 6 and the processed temperature data is shown in Table 7. As the same way as above mentioned, the curves of the raw data and the processed data are shown in Fig. 14. The curve of raw data is represented in blue and the curve of processed data is represented in red. It is obvious that the red curve removes the catastrophe point and make the curve smoother.

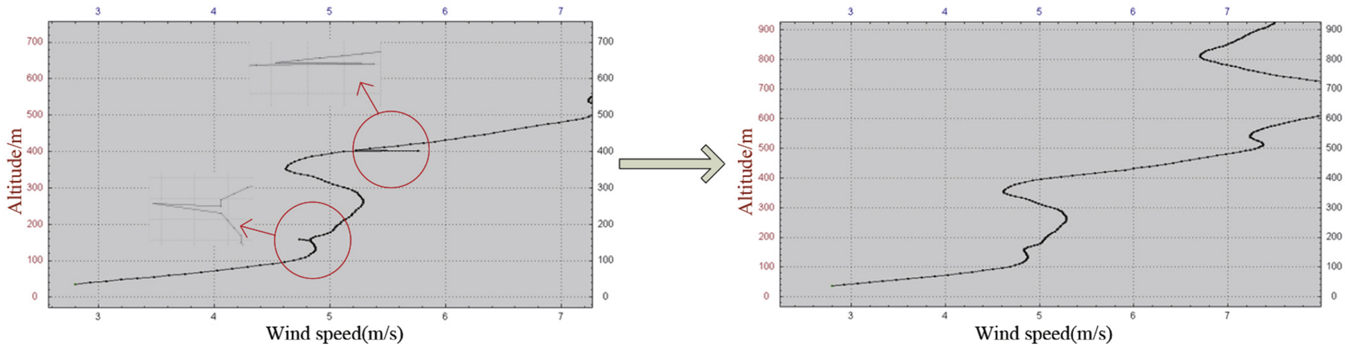


Fig. 13. The curve of wind speed data before and after processing.

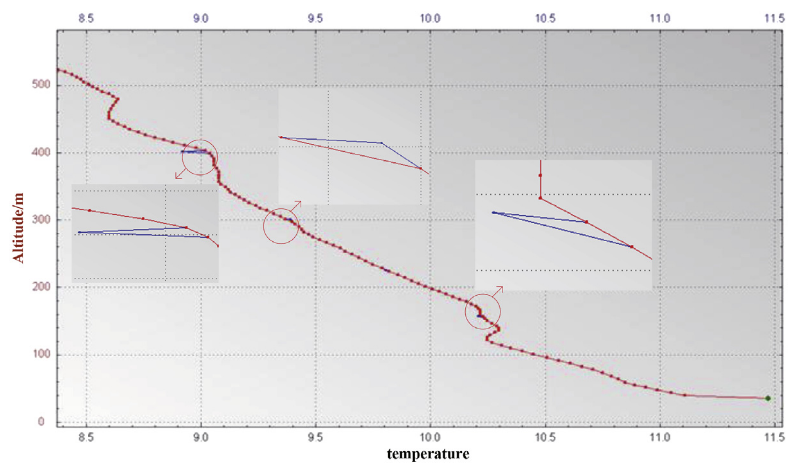


Fig. 14. Comparison of the temperature data before and after processing.

5. Conclusion

With the rapid development of the IoT, the application of IoT in the ocean will get a high degree of concern.

In this paper, we build a simple model of Cloud Computing and IoT with data interaction of multi-sensors and transmission to the cloud in the marine field. This simple model can represent the relationship of each sensor and the way of transmitting data information to the receiving terminal and the cloud. According to the big data from various kinds of sensors, we propose a data processing algorithm that can remove the outliers and add interpolation while the data distribution is loose. In order to verify the algorithm, we discuss the perspective of data processing algorithm and verify the feasibility of the proposed algorithm by comparing the image of raw data and the processed data. And the experiments show that the algorithm can improve the accuracy of data. Based on the processed data, we propose a mathematical model of route planning. The model can make the ship go to the destination more effective and plays a important role in the route layout. This proposed model can solve most situations in actual measurement except some critical situations.

The research work of this paper provides algorithm in marine data processing and a model in route planning, but there is still a lot of room for improvement. For example, the data processing algorithm is a basic method and the mathematical model of route planning is a hypothesis that ignore some critical situations. In the future, we will improve the data processing algorithm in ocean big data and add more influence factor and critical situations into the model, and gradually improve our research work.

Acknowledgment

This research is partially supported by Joint Fund of pre-Research on Equipment from Education Department of China (No. 6141A02022336). Research on multi-source data fusion and efficient processing of Marine information.

References

[1] S. Marston, Z. Li, S. Bandyopadhyay, A. Ghalsasi, Cloud computing the business perspective, Decis. Support Syst. 51 (1) (2011) 176–189.

[2] C. Chen, D. Chen, Y. Yan, G. Zhang, Q. Zhou, R. Zhou, Integration of numerical model and cloud computing, Future Gener. Comput. Syst. (2017).

[3] B. Varghese, R. Buyya, Next generation cloud computing: New trends and research directions, Future Gener. Comput. Syst. (2017).

[4] K. Kumar, Y.H. Lu, Cloud computing for mobile users: Can offloading computation save energy, Computer 43 (4) (2010) 51–56.

[5] L. Boloni, D. Turgut, Value of information based scheduling of cloud computing resources, Future Gener. Comput. Syst. 71 (2016).

[6] L. Wang, G.V. Laszewski, A. Younge, X. He, M. Kunze, J. Tao, C. Fu, Cloud computing: a perspective study, New Gener. Comput. 28 (2) (2010) 137–146.

[7] Q. Huang, Y. Yang, M. Shen, Secure and efficient data collaboration with hierarchical attribute-based encryption in cloud computing, Future Gener. Comput. Syst. (2016).

[8] C. Stergiou, K.E. Psannis, B.G. Kim, B. Gupta, Secure integration of iot and cloud computing, Future Gener. Comput. Syst. (2016).

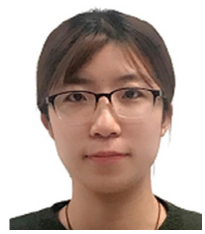
[9] F. Tao, Y. Cheng, L.D. Xu, L. Zhang, B.H. Li, Cciot-cmfg: Cloud computing and internet of things-based cloud manufacturing service system, IEEE Trans. Ind. Inf. 10 (2) (2014) 1435–1442.

[10] D. Miorandi, S. Sicari, F.D. Pellegrini, I. Chlamtac, Internet of things, Ad Hoc Netw. 10 (7) (2012) 1497–1516.

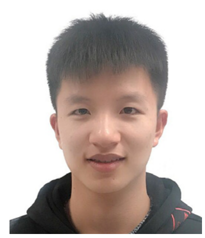
[11] M. Conti, A. Dehghantanha, K. Franke, S. Watson, Internet of things security and forensics: Challenges and opportunities, Future Gener. Comput. Syst. (2017).

[12] M. Ebrahimi, E. Shafieibavani, R.K. Wong, S. Fong, J. Fiaidhi, An adaptive meta-heuristic search for the internet of things, Future Gener. Comput. Syst. (2015).
[13] D. Miorandi, S. Sicari, F.D. Pellegrini, I. Chlamtac, Internet of things: Vision, applications and research challenges, Ad Hoc Netw. 10 (7) (2012) 1497–1516.
[14] T. Shon, J. Cho, K. Han, H. Choi, Toward advanced mobile cloud computing for the internet of things: Current issues and future direction, Mob. Netw. Appl. 19 (3) (2014) 404–413.
[15] J.A. Guerrero Ibanez, S. Zeadally, J. Contreras Castillo, Integration challenges of intelligent transportation systems with connected vehicle, cloud computing, and internet of things technologies, IEEE Wirel. Commun. 22 (6) (2015) 122–128.
[16] N.G. Wright, H.K. Chan, Low-cost internet of things ocean observation, in: Oceans, 2016, pp. 1–5.
[17] J. Yang, B. Jiang, B. Li, K. Tian, Z. Lv, A fast image retrieval method designed for network big data, IEEE Trans. Ind. Inf. PP (99) (2017) 1–1.
[18] Q. Zhang, Y. Wang, H.D. Han, Z.Y. Zhang, Development of internet of things in ocean environment monitoring, Ocean Technol. (2012).
[19] Y. Zhang, L. Yang, The mobile internet of things on the ocean based on digital shortwave communication, Sci. J. Control Eng. 3 (3) (2013).
[20] K. Hu, Z. Guo, Z. Che, F. Shen, Composition model of complex virtual instrument for ocean observing, J. Softw. 9 (5) (2014).
[21] G. Mulligan, The internet of things: Here now and coming soon, IEEE Internet Comput. 14 (1) (2010) 35–36.
[22] M.A. Feki, F. Kawsar, M. Boussard, L. Trappeniers, The internet of things: The next technological revolution, Computer 46 (2) (2013) 24–25.
[23] M.J. Fennessy, Dynamic environment charting - the next stage for marine electronic navigation, J. Navig. 52 (3) (1999) 367–377.
[24] J. Wozencraft, D. Millar, Airborne lidar and integrated technologies for coastal mapping and nautical charting, Mar. Technol. Soc. J. 39 (3) (2005) 27–35.
[25] B. Evans, J. Smith, J. Miller, Collaborative nautical charting and scientific seabed mapping, Sea Technol. 45 (6) (2004) 14–+.
[26] H. Lanziner, The genesis of electronic charting, Mar. Technol. Soc. J. 49 (6) (2015) 110–122.
[27] R. Plets, J. Dix, R. Bates, Marine geophysics data acquisition, processing and interpretation, Renew. Energy 60 (4) (2013) 269–283.
[28] E. Riedel, C. Faloutsos, G.A. Gibson, D. Nagle, Active disks for large-scale data processing, Computer 34 (6) (2001) 68–74.
[29] D. Yamaguchi, G.D. Li, M. Nagai, A grey-based rough approximation model for interval data processing, Inform. Sci. 177 (21) (2007) 4727–4744.
[30] R. Bose, J. Frew, Lineage retrieval for scientific data processing: A survey, ACM Comput. Surv. 37 (1) (2005) 1–28.
[31] M. Hammer, W.G. Howe, V.J. Kruskal, I. Wladawsky, A very high level programming language for data processing applications, Commun. ACM 20 (11) (1977) 832–840.
[32] M.R. Mosavi, Gps receivers timing data processing using neural networks: optimal estimation and errors modeling, Int. J. Neural Syst. 17 (5) (2007) 383.
[33] M. Basin, J.J. Maldonado, O. Zendejo, Mean-square state and parameter estimation for stochastic linear systems with Gaussian and Poisson noises, Int. J. Gen. Syst. 45 (5) (2016) 1–14.
[34] L. Liu, C.L.P. Chen, Y. Zhou, X. You, A new weighted mean filter with a two-phase detector for removing impulse noise, Inform. Sci. 315 (C) (2015) 1–16.
[35] L. Liu, Computing infrastructure for big data processing, Front. Comput. Sci. 7 (2) (2013) 165–170.
[36] S. Gan, S. Wang, Y. Chen, X. Chen, K. Xiang, Separation of simultaneous sources using a structural-oriented median filter in the flattened dimension, Comput. Geosci. 86 (C) (2016) 46–54.
[37] Q. Yang, N. Ahuja, K.H. Tan, Constant time median and bilateral filtering, Int. J. Comput. Vis. 112 (3) (2015) 307–318.
[38] P. Wielgosz, M. Krukowska, J. Paziewski, A. Krypiakgregorczyk, K. Stepniak, J. Kaplon, J. Sierny, T. Hadas, J. Bosy, Performance of ZTD models derived in near real-time from GBAS and meteorological data in GPS fast-static positioning, Meas. Sci. Technol. 24 (24) (2013) 5802.
[39] R.F. Leandro, M.C. Santos, R.B. Langley, Analyzing GNSS data in precise point positioning software, Gps Solut. 15 (1) (2011) 1–13.
[40] H. Wang, L. Wu, H. Chai, Y. Xiao, H. Hsu, Y. Wang, Characteristics of marine gravity anomaly reference maps and accuracy analysis of gravity matching-aided navigation, Sensors 17 (8) (2017) 1851.
[41] D.K. Schrader, B. Min, E.T. Matson, J.E. Dietz, Real-time averaging of position data from multiple GPS receivers, Measurement 90 (2016) 329–337.

Jiachen Yang received the M.S. and Ph.D. degrees in communication and information engineering from Tianjin University, Tianjin, China, in 2005 and 2009, respectively. He is currently a professor at Tianjin University. He was a Visiting Scholar with the Department of Computer Science, School of Science, Loughborough University, U.K. His research interests include data processing, multimedia processing, cloud computing and Internet of Things.



Chang Wang received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2017. She is currently pursuing the M.S. degree at the Tianjin International Engineering Institute, Tianjin University, Tianjin, China. Her research interests include marine data processing, cloud computing and Internet of Things.



Qiming Zhao received the B.S. degree in communication engineering from Tianjin University, Tianjin, China, in 2017. He is currently pursuing the M.S. degree at the Tianjin International Engineering Institute, Tianjin University, Tianjin, China. His research interests include marine data processing, cloud computing and Internet of Things.

Bin Jiang received the B.S. and M.S. degree in communication and information engineering from Tianjin University, Tianjin, China, in 2013 and 2016. He is currently pursuing the Ph.D. degree at the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. His research interests include data processing, multimedia processing, cloud computing and Internet of Things.

Zhihan Lv received his Ph.D. degree from the Ocean University of China in 2012. He worked in CNRS (France) as Research Engineer, Umea University (Sweden) as Postdoc Fellow, Fundacion FIVAN (Spain) as Experience Researcher. He was a Marie Curie Fellow in European Union's Seventh Framework Programme LANPERCEPT. He has also been an assistant professor in Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China since 2012. His research mainly focuses on Virtual Reality, Augmented Reality, Multimedia, Computer Vision, 3D Visualization, Graphics, Serious Game, HumanComputer Interaction, Networks, Bigdata, Software Engineering.

Arun Kumar Sangaiah received his Doctor of Philosophy (Ph.D.) degree in Computer Science and Engineering from the VIT University, Vellore, India. He is presently working as an Associate Professor in School of Computer Science and Engineering, VIT University, India. His area of interest includes software engineering, computational intelligence, wireless networks, bio-informatics, and embedded systems.

Please cite this article in press as: J. Yang, et al., Marine surveying and mapping system based on Cloud Computing and Internet of Things, Future Generation Computer Systems (2018), https://doi.org/10.1016/j.future.2018.02.032.