Implementation of voyage data recording device using a digital forensics-based hash algorithm

Ki-Taek Seong[1], Gwan-Hyung Kim[2]
[1]Department of Information Security, Tongmyong University, South Korea
[2]Department of Computer Engineering, Tongmyong University, South Korea

Article Info

Article history:

Received Apr 19, 2019
Revised Jul 5, 2019
Accepted Jul 17, 2019

Keywords:

Authenticity
Digital forensics
Hash algorithm
Integrity
NMEA 2000
Voyage data recorder

ABSTRACT

Identifying the causes of marine accidents is difficult because of problems in scene preservation, reenactment, and procuring of witnesses. Thanks to new regulations, larger vessels are now required to carry voyage data recorders (VDRs) and automatic identification systems (AISs). However, the content of these devices, which is created, stored, and managed digitally, has security vulnerabilities such as the potential for data modification. Therefore, when managing digital records it is important to guarantee reliability. To this end, we suggest a digital forensics-based digital records migration method using a hash algorithm to guarantee the integrity and authenticity of digital records.

Corresponding Author:

Ki-Taek Seong,
Department of Information Security,
Tongmyong University. South Korea.
Email: taichiboy1@gmail.com

1. INTRODUCTION

It is often difficult to identify the causes of marine incidents, not only because is it difficult to secure evidence associated with the location and case, but also because the intangible factors related to human resources, such as mariners, and the physical factors related to the environment, such as vessels, harbor facilities, and operational equipment, are intertwined.

To reduce these difficulties, recently built ships are equipped with and operated by digital devices encompassing the autopilot system, the voyage data recorder (VDR), the automatic identification system (AIS), and CCTV. Since marine traffic systems have become more complicated, marine incidents including collisions and sinking have increased; however, because evidence has been difficult to secure, making it difficult to identify the causes of incidents, the International Maritime Organization (IMO) under the UN has imposed a new regulation requiring the installation of a VDR and an AIS on all vessels [1]. The VDR was developed to store the data of the onboard digital devices to help in identifying the cause of a marine incident. The VDR data include the ship’s location, velocity, and heading as well as the conversations on board, the VHF radio communications on the ship’s operation, and the information displayed on the radar [2]. Since the stored electronic records are digital data, they are vulnerable to integrity and authenticity damage by the physical deterioration of storage devices or logic changes in the electronic recording system. The stored data should be protected from unauthorized access and from deletion by careless management [3]. This paper describes a method for preventing arbitrary deletion or alteration of data in the storage device by implementing digital forensics when recording voyage data in the storage device. In addition, this paper describes the implementation of digital forensics for identifying and characterizing the types of data that can be used as evidence for the investigation and assessment of marine incidents as well as in collecting, restoring, and analyzing the pertinent data.

2. RESEARCH METHOD

2.1. NMEA 2000–based data recording device

The integration of shipboard systems, sensing and control within systems, sharing of information, and collecting of data are occurring at an increasing rate on board vessels. A general shipboard configuration is shown in Figure 1. As shown in Figure 1, basic devices including sensors, actuators, and compass are linked to the conventional NMEA-0183 network, while other devices including sensors, actuators, engine, fuel, radar, GPS, autopilot, and ECDIS are linked to the instrument networks (NMEA 2000) [4, 5].

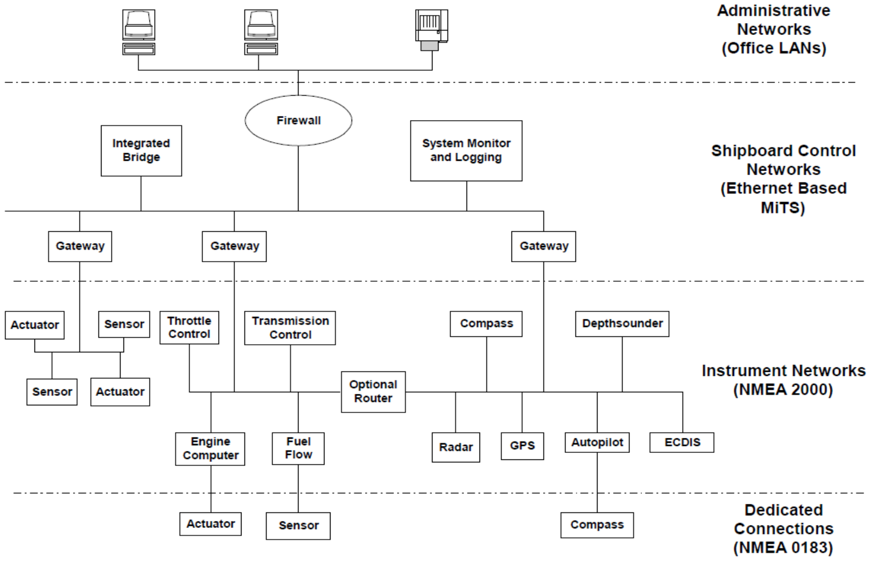


Figure 1. Shipboard networks and interfaces

2.2. Voyage data recording

The voyage data recorder (VDR) is intended for use in all passenger ships as well as in newly constructed cargo ships of 3000 gross tonnage and up. The S-VDR (simplified VDR) is for existing cargo vessels of the same category with a phase-in schedule requiring installation first in vessels of 20,000 gross tonnage and up, to be followed by others of 3000 gross tonnage and up. The S-VDR is not required to store the same level of detail as the VDR but should maintain a store, in a secure and retrievable form, of information concerning the position, movement, physical status, command, and control of a ship over the period leading up to and following an incident. Table 1 summarizes the two types of data items to be saved on the VDR [6-11].

Table 1. Data items to be recorded

No.	Data to be recorded	VDR	S-VDR
		Interface	
1	Date and time	IEC 61162	
2	Ship's position and datum used	IEC 61162	
3	Speed	IEC 61162	
4	Heading (from compass)	IEC 61162 Step, Synchro, Analog	
5	Bridge audio (by one or more microphones)	Audio	
6	Communications audio	Audio	
7	Radar, post-display selection	R, G, B, H, V	R, G, B, H, V (if available)
8	Depth	NMEA / IEC 61162	IEC 61162 interface (if available)
9	Main alarms	NMEA / IEC 61162, Contact, Analog	"
10	Rudder order and response	IEC 61162, Contact, Analog	"
11	Engine order and response	IEC 61162, Contact, Analog	"
12	Hull openings status	IEC 61162, Contact	"
13	Watertight and fire door status	IEC 61162, Contact	"
14	Accelerations and hull stresses	IEC 61162, Contact, Analog	"
15	Wind speed and direction	IEC 61162, Analog	"
16	AIS information		IEC 61162-2

The interfaces for the items that are saved need to satisfy International Electrotechnical Commission standard IEC 61162. The IEC 61162-1 standard refers to NMEA 0183, and IEC 61162-2 refers to NMEA 2000. Therefore, the VDR needs to satisfy the interface standard of NMEA 2000, or IEC 61162-2.

3. FORENSICS-BASED DATA RECORDING TECHNIQUES

3.1. Forensics-based accident investigation process

Because of the unique characteristics of marine accidents, the installation of data recording devices such as black boxes on airplanes is mandatory. However, such digital data are vulnerable to modification as they are restored and used within the device provided by the data recording device manufacturer. For the recorded data to have legal effectiveness, forensic techniques should be applied, and a process for submitting such data as legally binding evidence in the event of a marine accident has been proposed. As shown in Figure 2, data saved on VDR are saved using a digital forensics method, and the process of preserving the evidence happens when the ship starts to sink. Evidence preservation refers to the prevention of any manipulation of the saved data, which is an essential factor for the data to be considered as objective evidential data. Despite such evidence preservation, the data can still be manipulated through the interface, and thus it is necessary to confirm whether the data are from the time of accident.

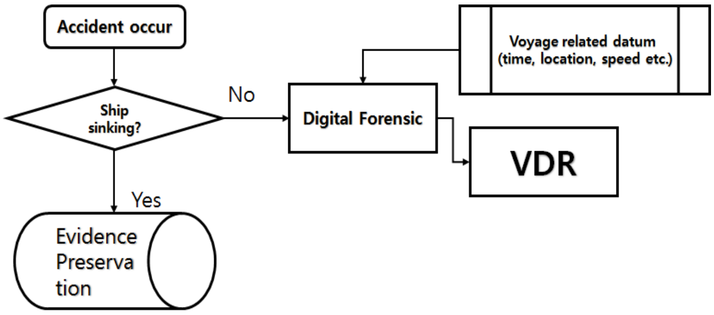


Figure 2. Process for investigation of marine casualties and incidents using digital forensic techniques

3.2. Forensics-Based Data Recording and Reproduction Techniques

Recent digital devices used in ships adopt the standard network, and the representative model is the NMEA 2000. Through this network, the devices transmit the voyage-related data to the NMEA 2000 network, which are then saved by the VDR, centering around the data summarized in Table 1. As mentioned before, the data need to be saved without any distortion and must be able to be verified as the data associated with a given period. Figure 3 shows the data recording process proposed in this paper.

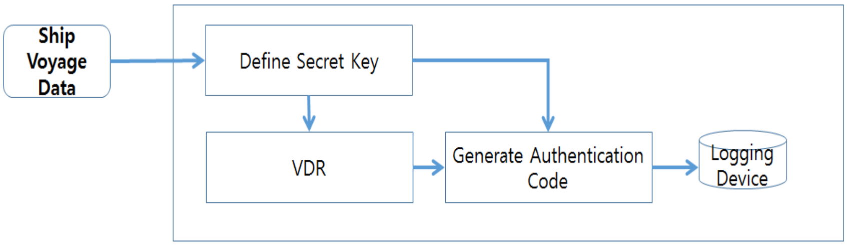


Figure 3. Forensics-based data recording process

In general, the voyage data are transmitted and saved transparently in the VDR. As shown in Figure 3, the proposed method receives the defined secret key from each data item and uses the data saved in the VDR and defined secret key to generate an authentication code, which is then saved on the logging device. The authentication code is a unique value for the saved data and is used to verify the authenticity of the data when manipulation takes place within the VDR. Figure 4 shows the authentication process for the data saved in the VDR using the logging device.

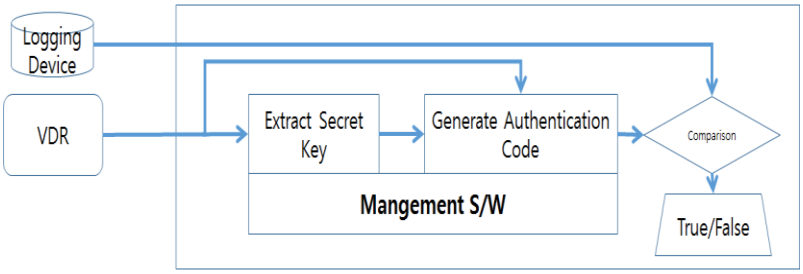


Figure 4. Forensics-based VDR data authentication process

Figure 4 shows the data authentication process using the logging device and VDR data from the forensics-based method. The authentication code, generated when the data was saved, is saved on the logging device. Then, when the new secret key extracted from the received VDR matches the authentication value as well as the data saved in the logging data, the integrity of the VDR data is considered confirmed, and when the information does not match, the data are considered to be damaged [12-27].

4. IMPLEMENTATION AND RESULTS

4.1. Test Bed

Using the method proposed in this paper, the network, as shown in Figure 5, was designed with the shipboard environment in mind. The GPS receiver, shown in Figure 5, receives the signal from the GPS and provides a variety of data including the two-dimensional location of the ship, its speed, its velocity, and the time. The NMEA 2000 is responsible for data communications, using a receiver and transmitter that satisfy the standards of the CAN 2.0 communication protocol, and the CPU manages the installed hardware and implements programs. In this arrangement, first the GPS receiver transmits the GPS signal, parsing it and changing it to the NMEA 2000 format, and then the logger uses this message to generate the secret key using the method given in Section 4.2 below, creating the authentication key. The voyage data are saved in the VDR, while the authentication key is saved in the logging device. Table 2 shows the specifications for the test bed. The actual test bed is shown in Figure 6.

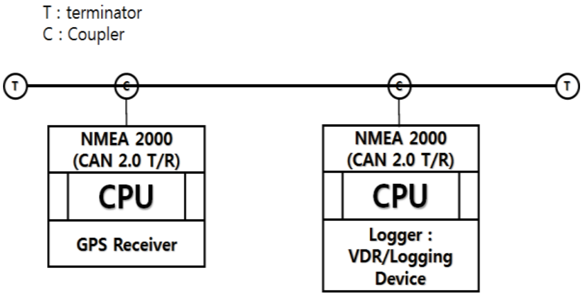


Figure 5. NMEA 2000 network test bed

Table 2. Specifications for test bed

GPS receiver		Logger	
Board name	Main role (Related hardware)	Board name	Main role (Related hardware)
CPU board	Central processing unit (Arduino UNO): External hardware control GPS message parsing Voyage data transmission	CPU board	Central processing unit (Arduino UNO): External hardware control Secret key extraction and authentication code generation, saved in logging device Voyage data saved in VDR
GPS board	GPS signal processing unit (Arduino GPS Shield): GPS signal receiving	Data logger	Saving unit (Arduino SD Shield): VDR and logging device role
CAN comm. board	Communication unit (Arduino CAN Shield): CAN 2.0 communication transmission processing	CAN comm. board	Communication unit (Arduino CAN Shield): CAN 2.0 communication transmission processing

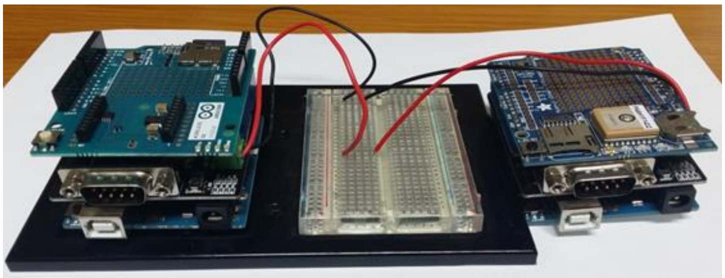
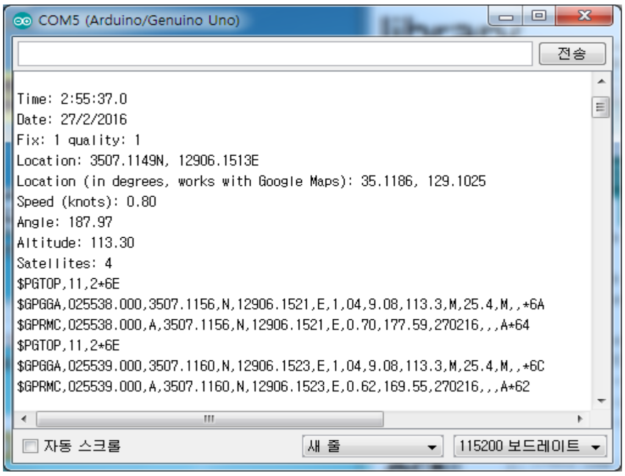


Figure 6. Test bed

4.2. Secret key and generation of authentication code

The information that can be received from this GPS receiver is as follows: time, data, location, speed (knots), angle, altitude, satellites, $PGTOP data, $GPGGA data, and $GPRMC data. Such information must be expressed in NMEA 0183 format in order to be transmitted through the NMEA 2000 Network. Figure 7 shows the actual NMEA 0183 format for receiving data.



```
Time: 2:55:37.0
Date: 27/2/2016
Fix: 1 quality: 1
Location: 3507.1149N, 12906.1513E
Location (in degrees, works with Google Maps): 35.1186, 129.1025
Speed (knots): 0.80
Angle: 187.97
Altitude: 113.30
Satellites: 4
$PGTOP,11,2*6E
$GPGGA,025538.000,3507.1156,N,12906.1521,E,1,04,9.08,113.3,M,25.4,M,,*6A
$GPRMC,025538.000,A,3507.1156,N,12906.1521,E,0.70,177.59,270216,,,A*64
$PGTOP,11,2*6E
$GPGGA,025539.000,3507.1160,N,12906.1523,E,1,04,9.08,113.3,M,25.4,M,,*6C
$GPRMC,025539.000,A,3507.1160,N,12906.1523,E,0.62,169.55,270216,,,A*62
```

□ 자동 스크롤 | 새 줄 | 115200 보드레이트

Figure 7. Data as parsed for NMEA 0183 format

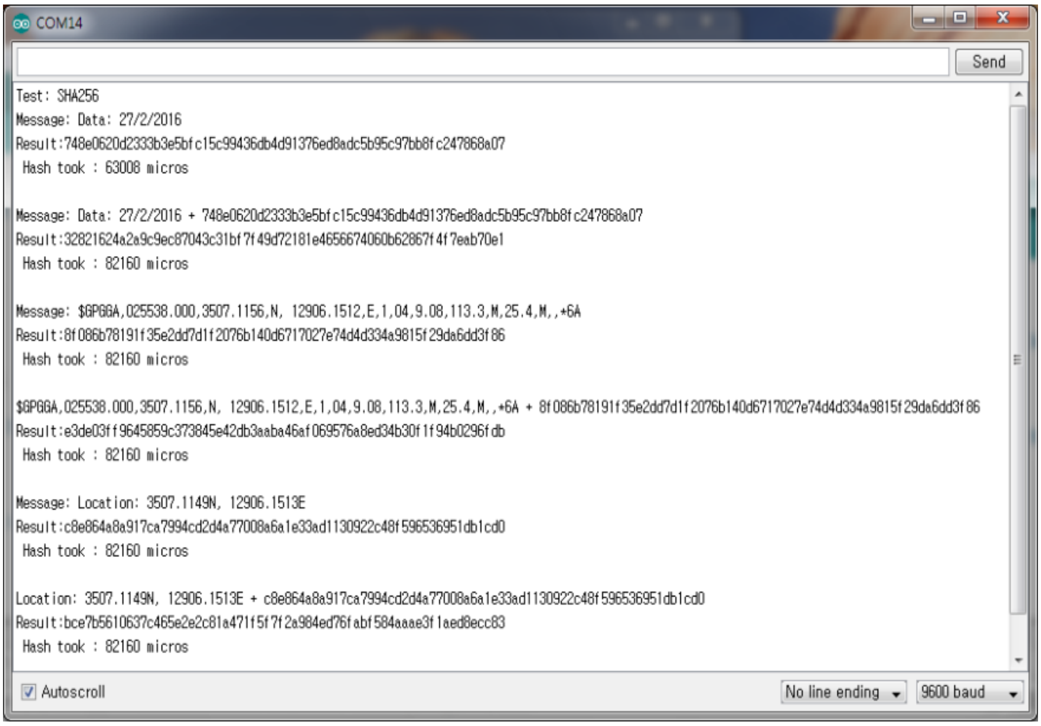
As all of the information received from the receiver is sting-type data, no distinguishing is needed when generating the secret key. Therefore, in this study, the GPS receiver generated the secret key and authentication code only for the three messages (time, location, and $GPRMC) that fall under Nos. 1 and 2 in Table 1. GPRMC stands for "Global Positioning Recommended Minimum Specific GNSS Data," and the $GPRMC message provides such information as UTC time/date, status, latitude/longitude, E/W/S/N indicator, speed over ground, and course over ground. The following process shows how the authentication code (Aunth_Code) is generated from the received message (Msg). First, the secret key is extracted from the message to be saved. Then, the authentication code is generated by combining this with the secret key.

Secret Key = Function(Msg, variables) ;

Aunth_Code = Function(Msg + Secret Key) ;

There are various methods for generating the secret key, including encoding the set value or using the string in a certain location. However, in order to enhance the security, it is best for it to be received from the respective message, and it is better for it to be of a fixed length regardless of the length of the message to be processed. The hash function, which is widely used to verify integrity because of its characteristic of creating a value of a given length, was used in this study. The security level was enhanced by using the hash function in generating both the authentication code and the secret key. There are various algorithms that implement the hash function, including MD5 (Message-Digest algorithm), SHA (Secure Hash Algorithm)–1/256, and SHA-256/224/512/383. MD5 is used to verify the original copy of a file or program because of its 32-bit–based fast processing speed. SHA-1 increase the hash value in order to enhance the security level, but it is known to be vulnerable to existing methods of attack. Therefore, the number of types of calculation and the hash value were increased, and SHA-256/224/512/383 were suggested.

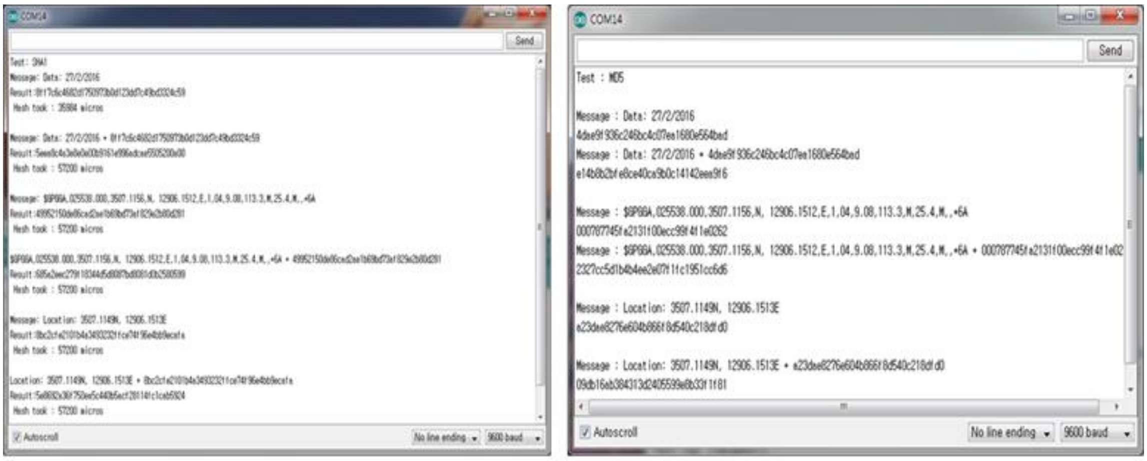
As shown in Figure 8, the secret key generated using the message “Date: 27/2/2016” was “0x748e06…”, and when this was combined with “Date: 27/2/2016” through hashing, the authentication code “0x32821…” was generated. Then the message was saved in the VDR, and the authentication key was saved in the logging device. Once the VDR message is modified, the authentication code in the authentication process will not match; therefore authenticity is guaranteed in the event that a match is obtained. Figure 9 shows the process for generating an authentication code that uses the SHA-1 and MD5 algorithms.



```
Test: SHA256
Message: Data: 27/2/2016
Result:748e0620d2333b3e5bfc15c99436db4d91376ed8adc5b95c97bb8fc247868a07
 Hash took : 63008 micros

Message: Data: 27/2/2016 + 748e0620d2333b3e5bfc15c99436db4d91376ed8adc5b95c97bb8fc247868a07
Result:32821624a2a9c9ec87043c31bf7f49d72181e4656674060b62867f4f7eab70e1
 Hash took : 82160 micros

Message: $GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A
Result:8f086b78191f35e2dd7d1f2076b140d6717027e74d4d334a9815f29da6dd3f86
 Hash took : 82160 micros

$GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A + 8f086b78191f35e2dd7d1f2076b140d6717027e74d4d334a9815f29da6dd3f86
Result:e3de03ff9645859c373845e42db3aaba46af069576a8ed34b30f1f94b0296fdb
 Hash took : 82160 micros

Message: Location: 3507.1149N, 12906.1513E
Result:c8e864a8a917ca7994cd2d4a77008a6a1e33ad1130922c48f596536951db1cd0
 Hash took : 82160 micros

Location: 3507.1149N, 12906.1513E + c8e864a8a917ca7994cd2d4a77008a6a1e33ad1130922c48f596536951db1cd0
Result:bce7b5610637c465e2e2c81a471f5f7f2a984ed76fabf584aaae3f1aed8ecc83
 Hash took : 82160 micros
```

☑ Autoscroll   No line ending   9600 baud

Figure 8. Generating authentication code for SHA-256



```
Test: SHA1
Message: Data: 27/2/2016
Result:8f17c6c4682d1750973b0d123dd7c49bd0324c59
 Hash took : 35984 micros

Message: Data: 27/2/2016 + 8f17c6c4682d1750973b0d123dd7c49bd0324c59
Result:5eee8c4a3e0e0d00b9161e996adcae5505200a00
 Hash took : 57200 micros

Message: $GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A
Result:49952150de86cad2ae1b68bd73ef829a2b80d281
 Hash took : 57200 micros

$GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A + 49952150de86cad2ae1b68bd73ef829a2b80d281
Result:685a2eec279f18344d5d8087bd8081d0b2580599
 Hash took : 57200 micros

Message: Location: 3507.1149N, 12906.1513E
Result:8bc2cfe2101b4a34932321fce74f96e4bb9ecafa
 Hash took : 57200 micros

Location: 3507.1149N, 12906.1513E + 8bc2cfe2101b4a34932321fce74f96e4bb9ecafa
Result:5e8692a36f750ee5c440b5ecf28114fc1ceb5924
 Hash took : 57200 micros
```

☑ Autoscroll   No line ending   9600 baud

(a)

```
Test : MD5

Message : Data: 27/2/2016
4dae9f936c246bc4c07ea1680e564bad
Message : Data: 27/2/2016 + 4dae9f936c246bc4c07ea1680e564bad
e14b8b2bfe8ce40ca9b0c14142eea9f6

Message : $GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A
000787745fa2131f00ecc99f4f1e0262
Message : $GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,+6A + 000787745fa2131f00ecc99f4f1e02
2327cc5d1b4b4ee2e07f11c1951cc6d6

Message : Location: 3507.1149N, 12906.1513E
a23dae8276e604b866f8d540c218dfd0

Message : Location: 3507.1149N, 12906.1513E + a23dae8276e604b866f8d540c218dfd0
09db16ab384313d2405599e8b33f1f81
```

☑ Autoscroll   No line ending   9600 baud

(b)

Figure 9. Generating authentication code for SHA-1 (a) and MD5 (b)

Table 3 summarizes the test results. The security of the hash algorithm shows the frequency of clashing, in which the same output value occurs for different input values. The lower the frequency, the higher the security level, and thus the SHA-256 method is the most stable, but it requires a longer processing time, as shown in Table 3.

Table 3. Summary of results

Algorithm (Output bits)	Input	Key value / Execution time us	Aunth_Code value / Execution time us	Security
MD5	String 1	0x4dae9… / 1504	0xe14b8… / 13,520	Low
(128)	String 2	0x00078… / 13,524	0x2327c… / 13,524	
	String 3	0xa23da… / 13,524	0x09db1… / 13,524	
SHA-1	String 1	0x8ff7c6… / 35,984	0x5eea8… / 57,200	Medium
(160)	String 2	0x49952… / 57,200	0x685a2… / 57,200	
	String 3	0x8bc2c… / 57,200	5e8692a… / 57,200	
SHA-256	String 1	0x748e0… / 63,008	0x3282162… / 82,160	High
(256)	String 2	0x8f086b… / 82,160	0x45d52… / 82,160	
	String 3	0xc8e86… / 82,160	0x7bb00… / 82,160	

Note:
String 1 = “Data: 27/2/2016”
String 2 = “ $GPGGA,025538.000,3507.1156,N, 12906.1512,E,1,04,9.08,113.3,M,25.4,M,,*6A”
String 3 = “Location: 3507.1149N, 12906.1513E”

5. CONCLUSION

In this study, a voyage data recording method using digital forensics was proposed and implemented. The message received from the GPS receiver is changed into the NMEA 0183 format through parsing. The time and location information for the VDR record item is used to generate a secret key; this is used to generate the authentication code for the message and is then saved. A hash function is used to generate a secret key for each of the saved messages. In generating the authentication code, the secret key and the message are combined, and the hash function is used in order to guarantee the unique characteristics of the message and its authenticity. In this way, the hash function was used both to enhance the level of security and to uniquely identify it. Three widely used types of hash algorithms, SHA-1, SHA-256, and MD5, were implemented.

The actual data saved in the VDR can be in string units or file units. Whereas irregular, short pieces of information are transmitted in the form of a message, analog signals, including voice recordings and radar images, can be saved in file units. The advantage of this method is that it can easily be applied to the traditional VDR since it can be applied to the aforementioned files. The use of forensics for VDR data to be used as evidential data in the event of marine accidents can enhance the objectivity of the evidence, and it is our hope that the proposed recording device will be widely used.

REFERENCES

[1] International Maritime Organization (IMO), “International Convention for the Safety of Life at Sea (SOLAS),” Chapter V, International Maritime Organization, London, 1974.

[2] Ministry of Land, “Transport and Maritime Affairs, Notice 2012-075: Vessel Equipment Standards,” Article 108.7 (Voyage Data Recorder), 2012.

[3] Ministry of Land, “Transport and Maritime Affairs, International Maritime Organization 55th Maritime Safety Committee Final Report,” pp. 29-32, 2009.

[4] J. H. Huh, et al., “Design of NMEA2000 CAN Bus Integrated Network System and Its Test Bed: Setting Up the PLC System in Between Bridge—Bow Room Section on a Container Ship as a Backbone System,” *Lecture Notes in Electrical Engineering*, Springer, vol. 2016, pp. 191-204, 2015.

[5] A. Ninomiya, “About the outline of NMEA 2000®,” Tokyo University of Mercantile Marine, pp. 52-54, 2011.

[6] S. C. Austin and P. A. Wilson, “Maritime voyage data recorder study by the European maritime data management project,” *International Journal of Maritime Engineering*, vol. 151, pp. 13-24, 2009.

[7] Uchijima, et al., “Development of VDR (Voyage Data Recorder),” *Nihon Musen Giho, JRC Review, JRC Authors*, vol. 67, pp. 42-45, 2016.

[8] Hsu, et al., “Constructing an Efficient State Space Query System for the Voyage Data Recorder,” *Frontiers in Artificial Intelligence and Applications, IOS Press*, vol. 2015, pp. 294-305, 2015.

[9] J. Kang, et al., “Development of Remote Alarm Module with Playback functions in Voyage Data Recorder,” *SICE -Annual Conference*, vol. 2009, pp. 3I06-4, 2009.

[10] C. Jung, et al., “The development of Ethernet based radar and ECDIS image processing for voyage data recorder,” *Proceedings of the 14th International Conference on Control, Automation and Systems, Gyeonggi-do, Korea, Republic of,* pp. 963-966, 2014.

[11] Maritime navigation and radiocommunication equipment and systems - Shipborne voyage data recorder (VDR) - Part1: Performance requirements, methods of testing and required test results, IEC 61966 - 1, May, 2013.

[12] Myeong H. B. and Sangjin L., “A new investigation methodology of marine casualties and incidents using digital forensic techniques,” *Journal of The Korea Institute of Information Security & Cryptology*, vol. 23, pp. 515-530, 2013.

[13] M. Piccinelli and P. Gubian, “Modern ships voyage data recorders: A forensics perspective on the Costa Concordia shipwreck,” *Digital Investigation*, vol. 10, pp. 41-49, 2013.

[14] A. Frieze and T. Johansson, “On the insertion time of random walk cuckoo hashing,” John Wiley & Sons, Ltd, vol. 54, pp. 721-729, 2019.

[15] Z. Han, et al., “A novel routing algorithm for IoT cloud based on hash offset tree,” *Future generations computer systems, Elsevier Science B.V., Amsterdam,* vol. 86, pp. 456-463, 2018.

[16] S. L. Garfinkel and M. McCarrin, “Hash-based carving: Searching medi a for complete files and file fragments with sector hashing and hashdb,” *Digit al Investigation,* vol. 14, pp. 95-105, 2015.

[17] M. G. Noblett, et al., “Recovering and Examining Computer Forensic Evidence,” *Journal in Forensic Science Communications*, vol. 2, pp. 1-13, 2000.

[18] W. G. Henrique, “Anti Forensics: Making computer forensics hard,” Code Breakers III, Sao Paulo, Brazil, 2006.

[19] M. C. Stamm and K. J. R. Liu, “Anti-forensics of digital image compression,” *IEEE Trans. Inf. Forensics Security,* vol. 6, pp. 1050-1065, 2011.

[20] M. C. Stamm, et al., “Anti-forensics of JPEG compression,” *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 1694-1697, 2010.

[21] M. C. Stamm and K. J. R. Liu, “Wavelet-based image compression anti-forensics,” *Proc. IEEE Int. Conf. Image Processing*, pp. 1737-1740, 2010.

[22] M. C. Stamm, et al., “Forensics vs. anti-forensics: A decision and game theoretic framework,” *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Kyoto, Japan,* pp. 1749-1752, 2012.

[23] M. Chen and W. Hu, “Research on BatSLAM Algorithm for UAV Based on Audio Perceptual Hash Closed-Loop Detection,” *International journal of pattern recognition and artificial intelligence*, vol. 33, 2019.

[24] G. Collom, et al., “Fast Mesh-to-Mesh Remaps Using Hash Algorithms,” *SIAM Journal on Scientific Computing,* vol. 40, pp. C450-C450, 2018.

[25] X. Wang and J. Yu, “Studies on An Online Density Sensitive Hash Algorithm,” *MINIMICRO SYSTEMS - SHENYANG,* vol. 39, pp. 1068-1073, 2018.

[26] Y. Qin, et al., “Multi-stage IPv6 Routing Lookup Algorithm Based on Hash Table and Multibit Trie,” *MINIMICRO SYSTEMS -SHENYANG,* vol. 39, pp. 893-898, 2018.

[27] M. A. Abdulhayoglu and B. Thijs, “Use of locality sensitive hashing (LSH) algorithm to match Web of Science and Scopus,” *Scientometrics, Akademiai Kiado Rt.,* vol. 116, pp. 1229-1245, 2018.