CHRIS/11/12.4 A

11th CHRIS MEETING
IHB, Monaco, 16-18 November 1999

TECHNOLOGY ASSESSMENT WORKING GROUP (TAWG)
(Mr. Michael CASEY)

Encryption and ENCs
The Technology and Policy Issues

EPG DRAFT Report

Version 2.0
November, 1999

Encryption and ENCs
The Technology and Policy Issues
EPG DRAFT Report
November, 1999

Table of Contents

1) Security Systems and ENCs

1.1) The IHO issue

Several countries producing ENCs have expressed concerns over the protection of ENC data for both legal liability and revenue protection reasons. Some HO’s feel that ENCs need to be protected from piracy and/or deliberate tampering. The IHO, in an effort to stabilize the introduction of the ECDIS technology, aims to provide as well-coordinated approach to the delivery of ENCs as practical. Such coordination would extend to the ways and means the data would be secured against tampering or piracy. There is a well founded concern that if many HO's introduce widely different security schemes, the impact upon the mariner will not enhance the introduction of ECDIS but be a set-back. In general "ease-of-use-promotes-use" is the prevailing sentiment today. Unless the system is user friendly and implemented in a well coordinated way encryption might prove to be a serious setback to the global acceptance of ECDIS.

The IHO CHRIS Committee tasked the Technology Assessment Working Group (TAWG) with investigating the issues surrounding data security systems and an Encryption Project Group (EPG) has been formed to carry out this study.

The Terms of Reference of the EPG state as their Objective

> "To assess the potential issues surrounding the encryption of ENC data, to examine the various existing encryption methods with respect to efficiency and practicability both at HOs and by the marine end user, and advise the TAWG accordingly through a report"

The EPG is formed of about 25 individuals who have expressed an interest in the subject. The work has been carried out by email correspondence and through the use of the Open ECDIS Forum. The EPG extends its thanks to SevenCs for supporting this Internet-based tool.

This paper has been prepared as a report to CHRIS 11, November, 1999.

There are two main issues involved in this subject: security technology, and, inherent security policy issues.

1.2) Why Encryption ?

For HOs and their agents there are several motivations for encryption. The following is a list of the 5 objectives of a security system.

The 5 Objectives Of A Security Schema

1. *protection against non-deliberate virus or other unintentional corruption of the data*
2. *demonstration to the end-user of the legitimacy and integrity of the data*
3. *demonstration of the ownership of the data and implied or explicit copyright protection*
4. *protection against deliberate corruption or manipulation of data*
5. *protection against data piracy*

Typically it is this last item that most people usually think of when they think of encryption. The protection of copyright data from acts of deliberate data piracy is a concern for agencies whose existence is dependent upon some form of cost recovery. Clearly the encryption of data products is a very explicit declaration of the seriousness attached to copyright protection. It is a clear and deliberate act which signifies the importance of data security along the distribution chain. Additionally any protection against deliberate tampering of the data would be a step forward in risk avoidance.

To the five objectives above one could add the following three which relate more to the implementation of the security system.

- *Implementation simplicity*
- *End-user simplicity*
- *Speed of implementation*

The system must be relatively easy to implement particularly if it is to be a global standard.

In a perfect world an encryption scheme would be practically unbreakable yet its effect on the end-user totally transparent. The latter is an important issue since all encryption schemes put some burden on the end user. The limits to what mariners are willing to put up with, in terms of the burden the decryption process takes, will, in the end, dictate the true level of security attained. A largely unbreakable scheme is possible but only when the mariner and everyone else in the distribution chain cooperate and agree to some fairly rigid protocols. We cannot always expect this to happen and so some compromises must be made to find the right level of security. As a general principle one requires a scheme that 1) costs as much to break as to legitimately purchase the data and 2) is transparent enough for traditional clients to accept.

Finally the system must be implemented within a relatively short time period.

A well designed security system will invoke the appropriate level of security with minimal burden placed upon the end-user.

2) Primer on Encryption and Authentication

Encryption is not a new technology. The first use of encryption dates back thousands of years. Over time the technology has changed as the encrypters try and stay one step ahead of the decrypters.

Historically the purposes of encryption have largely been for military or political reasons although in the latter part of the 20th century it has found a commercial home. Most recently the state of the art in encryption technology available to the general public has become so advanced as to be considered a security problem. Largely unbreakable codes for example can prevent police from carrying out legal investigative search techniques or can allow foreign states to access security technology that can be used against the nations that have developed them.

Data Authentication, on the other hand, although related, fulfills a narrower objective verification that the data set has arrived in the same state that it was released by the HO. Authentication therefore satisfies the first 4 objectives of security but not necessarily the protection against piracy.

A primer on the technology has been developed and is attached as Appendix 1.

A great deal has been written about the subject of encryption. The Internet is an easy source of up-to-date information about the science of encryption and the state-of-the-art of commercially available software.
(http://www.ssh.fi/tech/crypto/) is a good place to start.

Bruce Schneier is the author of “Applied Cryptography” one of the key textbooks in the field. His web site (http://www.counterpane.com/) offers essays on the various technical issues on encryption and includes free downloadable encryption routines such as BLOWFISH.

3) Impediments to Achieving The Goals Of A Security System

The major impediments to achieving the goals of a security system can be categorized as follows:

- **Weaknesses in the end-to-end system**
- **Lack of a standardized encryption methodology**
- **Type approval limitations**
- **IMO concurrence**
- **Lack of acceptance by end users**
- **Complexity of global Key Management**

3.1) Weaknesses in the end-to-end system

Encryption is frequently seen as a complete security solution. In reality it is only part. Data that is encrypted can still be compromised by a number or non-cryptoanalytic attacks. For example, you don't need to break the code to read an encrypted message. You can watch over the shoulder of the person doing the encoding, or steal decryption keys, or bribe someone, etc., etc. Physical, and other forms of electronic security are still necessary. Often the weakest encryption methodology is the strongest link in an overall security scheme as the more obvious loopholes are ignored..

In the case of ENC encryption, the data as well as the keys for the encrypted data are being made available to a number of parties in the distribution chain. Some will have access to ECDIS units as well (systems suppliers, service people etc.). In the end an ECDIS unit will have to be able to display the ENCs, independent of whether the data was at some point encrypted. At that point the security is vulnerable.

The ECDIS is thus, a weak point in the security system. There are other weak points in the system as well. The HOs that provide the data, and the personnel that work there will have access to the ENCs. As will RENCS and any dealers, agents or system manufactures that distribute the data. The data will have to be trusted to software systems, that could easily make copies of the clear form ENCs as they processed them.

3.2) Lack of a standardized encryption methodology

There is great value in having only one standard method for encryption so that users are not saddled with a variety of decryption schemes, multiples keys etc.

3.3) Type approval

No universal security system exists - yet several systems are in the final stages of type approval. A security system which incorporated decryption would require sufficient changes to the software to void the certification. Hence the system would have to undergo re-certification. The degree of re-testing would be restricted to those feature considered at risk. It is expected that the re-certification would be relatively straightforward.

3.4) IMO politics

The widespread acceptance of ECDIS in the marine community is dependent upon the continued support of the IMO. As encryption was not a feature of S-52 and its addition could not be considered a minor deviation from the original intent of the specifications, the IMO might pose a political problem if some nations decide to make it an issue..

3.5) Complexity of global Key Management

A pivotal issue whenever encryption is involved is Key Management. Decryption requires a key to unlock the data. Keys can be used for individual files or for a set. Typically one would expect that ENCs distributed on CD-ROMs would contain one or more ENCs that the user might want. Access management is a secondary role for encryption that allows the distributor to place more information on CD-ROMs that users might want. This might make for a more economically viable distribution method. If each file is encrypted uniquely then users can get just the files they have paid for. This would involve a key for each file. To control pirating the keys are uniquely linked to a specific ECDIS. Ships with more that one system therefore need separate keys for each. The management of these keys and their administrative requirements (such as license renewal dates) must be managed in a way that isn’t a burden on the end user.

3.6) Lack of acceptance by end users

The user community for ENCs have a limited tolerance for user-unfriendly features. This has been amply demonstrated through the various sea trials conducted around the world. Professional mariners want tools that are immediately helpful and provide unambiguous and useful information in a timely manner. Tools must be well designed and ergonomically structured to assist the mariner in conducting incident-free voyages. If the security system is complex for the user to understand and administratively difficult to manage we can expect a string resistance to acceptance. The consequences of this are unknown but likely to be painful. The security system should be largely transparent to the end-user.

4) General Options for Implementing A Security System

There are a number of options in considering a security system:

1. Firstly, one could choose to **do nothing**. A level of protection could be implied through the signing of license agreements stating the conditions of copyright.

2. Secondly, one could **shift the burden** of protection to the distribution chain through agent agreements.

3. Thirdly, one could **watermark** the data is some unique but hidden way in which copyright violators could be traced.

4. Fourth, one could implement a HO or **RENC specific encryption** scheme.

5. Fifth one could devise and implement **a globally acceptable encryption scheme** agreeable to all HO's and RENCs.

6. Finally one could **change the S57 standard** so that layers or S-57 objects can be optionally encrypted

Each approach has its advantages and disadvantages.

4.1) Do nothing

Doing nothing is always an option. In fact that is what a number of HO's are doing at the moment. Protection is assumed to follow from signed license agreements and willingness to litigate against pirates.

4.2) Shift The Burden

HO's can avoid the encryption issue by shifting the security burden to system manufacturers. In that case each manufacturer could design a system-specific scheme around their SENC. The onus would be placed on the manufacturer to provide proof of security and they could be made liable for piracy of data for systems which did not meet security performance specifications. The chief advantage is that there is no need for a universally acceptable method for encryption. A disadvantage is that users must be able to acquire their SENCs anywhere in the world, an additional burden on a systems manufacturer with a limited distribution chain. Internet-based distribution might alleviate this problem. A second disadvantage might be that users, tied to one data supplier, feel they were not benefiting from a more open and competitive environment.

4.3) Watermark The Data

Watermarking the data demonstrates that HO’s or RENCs are building in a tracing mechanism by which pirates can be successfully prosecuted for copyright violation.

4.4) Encrypt The Data

Whole file encryption is straightforward and does nothing to the existing standard. Encrypted data files are not S-57 until they are decrypted. The UK has developed a successful encryption scheme in its ARCS product as has C-Map and others. Whole file encryption is an immediate option. The disadvantage is that a variety of encryption scheme will be incompatible and make key management an unacceptable burden on the end user.

4.5) Add Encryption to S57

Encrypting layers or individual objects, on the other hand, is much more flexible. Using this methodology, decisions could be made about what to encrypt and what not to encrypt. For example, navigation-critical information, information that would be necessary to avoid maritime incidents, could be left unencrypted, while less critical information could be encrypted, so that only those who paid for the service could access them. The disadvantage with this method, is that it would require an overhaul of the existing S-57 standard. And it would certainly add complexity to S-57. Given the desire by HO’s to freeze the format until a significant critical mass of ENCs has been created means this approach is unworkable in the short run.

5) Stepping To The Side: Ethical Dilemmas and Possible Revenge Effects:

5.1) Data Security Vs. Navigation Safety

5.1.1) The Ethical Issue

The ethical issue in encryption is a simple one: can a HO or its Agent deny access to an ENC that the user has not paid for ? The problem is confounded by the fact that the data may in some cases be resident on the ship but in an encrypted form. It seems inappropriate for agencies whose mandate is the provision of information to enable safe navigation to deny access - even when they have a legitimate business reason for doing so. Since ECDIS equipped ships may not carry up-to-date paper charts, the ENC may be their sole access to a nautical chart.

5.1.2) The Legal Issue

A related issue is the legal one; is there a legal liability issue if encryption denies access ? Informal discussions with an admiralty lawyer indicated that as long as a substantial effort had been made to contact the licensee about license renewal and if the grace period was considered of reasonable length then the court would likely find that the licensee had sufficient time available to renew a license and knowingly decided against renewal.

5.1.3) License Periods, Warning Periods & Grace Periods

Encryption means privileged access and in a licensing environment it means privileged access for predetermined intervals. The situation can be summarized in the following graphic. Towards the end of the standard license period the users are warned of the expiry date. The warnings can be delivered in a variety of ways but are designed to remind users of the approaching end date of the license. This is referred to as the Warning Period. For example a period of two months prior to the expiry date, warning messages are given or a window on the display shows the end date or days-to-go. At the end of the Warning Period is E-Day for expiry day. This marks the beginning of the Grace Period. The Grace Period allows a level of service lower than the standard level and runs for a period yet to be determined. At the end of the Grace Period is the T-Day for termination day. After this date the data is unavailable until action is taken to renew the license.

Table 1: The Warning and Grace Periods

License Period		Out of License Period			
Standard	Warning Period	E	Grace Period	T	
	░	▒	▒	▓	█

5.1.4) The WEND Principle For Access Termination

Debate about termination strategy was conducted at the WEND meeting in Singapore in January 1999. The consensus of that meeting was that users who had previously licensed data would not be denied access to that same data despite the fact that the license period had expired.

5.2) Possible Revenge-Effects in Implementing Encryption

The history of the introduction of a new technology always has unintended effects even its proponents had not imagined. Sometimes these effects are beneficial, some are benign, some annoying and some dangerous. These latter occurrences have been dubbed technology "bite-backs" or "revenge effects" (Tenner, 1996). The technology itself doesn't provide the revenge effect - it is in the way that it changes people's perception or habits that things really start to go astray. The field of "Risk Homeostasis" is littered with examples of technologies that, when implemented, had unforeseen effects. An increase in the number of road accidents with cars equipped with anti-lock brakes has been attributed to the fact that the technology falsely raised the confidence level of drivers when operating under poor road conditions, leading to higher speeds and shorter car-to-car gaps. The technology worked as it was designed but the effect did not. When implementing technology the whole system has to be included and this includes the human element which is harder to model than any other component. What revenge effects might encryption introduce and what can we do to mitigate against them ?

5.2.1) Reliability problems leads to fewer new buyers:

Introducing encryption comes at a time when the market remains unconvinced about the merits of ECDIS. Any problems (real or perceived) caused by encryption might makes potential buyers hold off purchasing ECDIS - a negative consequence for marine safety overall. **The security system must be reliable.**

5.2.2) Users abuse the security and circulate unencrypted data:

Some users may gain access to the decrypted data copy it and make it available to others, much in the same way software is innocently pirated and passed around for try-before-you-buy purposes**. The security system must make pirating difficult and apparent to end users.**

5.2.3) Weak links in distribution chain:

Encryption requires the cooperation of all units in the distribution chain to abide by given protocols. An agent under pressure to quickly solve a problem and allows clients access a decryption key "just this once" (a relatively common practice in the software industry) can provide a serious fault in the encryption armor. **The security system must be administratively easy to manage and have a fast response.**

5.2.4) Users lean towards pirated data and this becomes acceptable behavior:

Users who gain access to pirated data will tend to rationalize their behavior and grow to accept the concept as acceptable leading to more widespread violations. **Users need to be educated to the overall benefits of a good security system.**

5.2.5) Encryption leads to loss in navigation safety due to lack of ENCs at crucial moment:

HOs exist to help prevent marine accidents. A loss of chart information at a critical juncture might lead to a calamitous incident the consequences of which are many orders of magnitude larger than the loss of licensing revenue. **The security system cannot deny access to nautical chart data users had previously purchased.**

6) EPG Opinions

In order to gauge the opinions of EPG members on the major implementation dilemmas, a poll was taken via a questionnaire. The questions were designed to firstly assess the perceived threat and secondly, the perceived threat response.

Figure 1 shows the data summarized statistically with the range of responses (the central vertical line from lowest to highest), the mode (the small error bar symbol showing the most frequently rated value) and the percentiles (the boxes going from 10% value to 90%). For example on the question of a "grace period" all response fell in the "almost certain" to "very likely" range with the "almost certain" being the most frequent answer. Contrast this with the spread of responses on the "termination strategy".

Range, Mode, 90 & 10 Percentiles

almost certain	5
very likely	4
likely	3
unlikely	2
very unlikely	1
almost certainly not	0

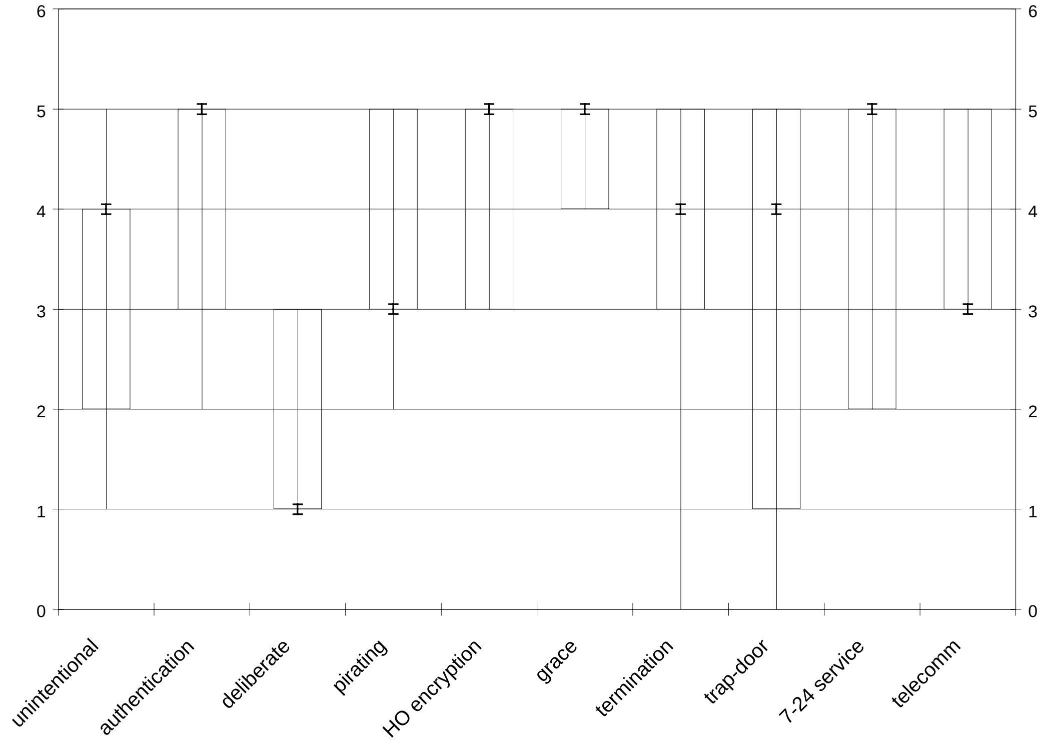


Figure 1 Range, Mode and 90/10 Percentiles

6.1) Meaning of the Results

1. **Non-deliberate virus or other invasive processes will unintentionally corrupt the data:** Most respondents feel that this threat is a likely occurrence.
2. **End-users require proof of the legitimacy and integrity of the data:** Authentication is seen by most as a way to identify corruption of data
3. **Deliberate corruption or manipulation of data will occur:** This is seen as a unlikely threat
4. **Pirating of data will occur on a scale sufficient to warrant some form of preventative action:** This seen as a likely occurrence
5. **Encryption will be used by many HOs and agents in the delivery of ENC and ENC updates:** Encryption is a most likely outcome
6. **A "grace period" will enable mariners access to data for some time beyond the legal or contractual date:** Everyone agrees that some form of "grace period" will allow access beyond the official license period.
7. **After a grace period access to the data would unavailable if licensing requirements were not met:** This is the "Termination Strategy" to be adopted by each HO. There is a wide range of perceived responses.
8. **Some form of emergency pass key or trap door would allow mariners in default access to sufficient data for safe operations into the next port of call:** This refers to the degree of "termination". A "trap door" is a way of providing some or all of the data even past the "grace period". Again there is a wide range of perceived responses.
9. **License renewal would be available on a 7 days a week/24 hours per day basis via communication link:** This refers to the likely service available to users to renew lapsing licenses. The responses indicate such a service is a likely (but not certain) event.
10. **All ships having ECDIS will have sufficient telecomm links to re-acquire license strings:** This refers to the telecomm capability of users to receive a renewed license key. The responses indicate this a likely (but not certain) event.

Figure 2 shows the perceived threat, simplified to show only the polar ends of the values. While deliberate tampering with the data appears unlikely, some form of authentication is needed to assure all participants in the chain that valid data has arrived at the ECDIS terminal. Piracy is seen as a meaningful threat.

Figure 2: The Threat

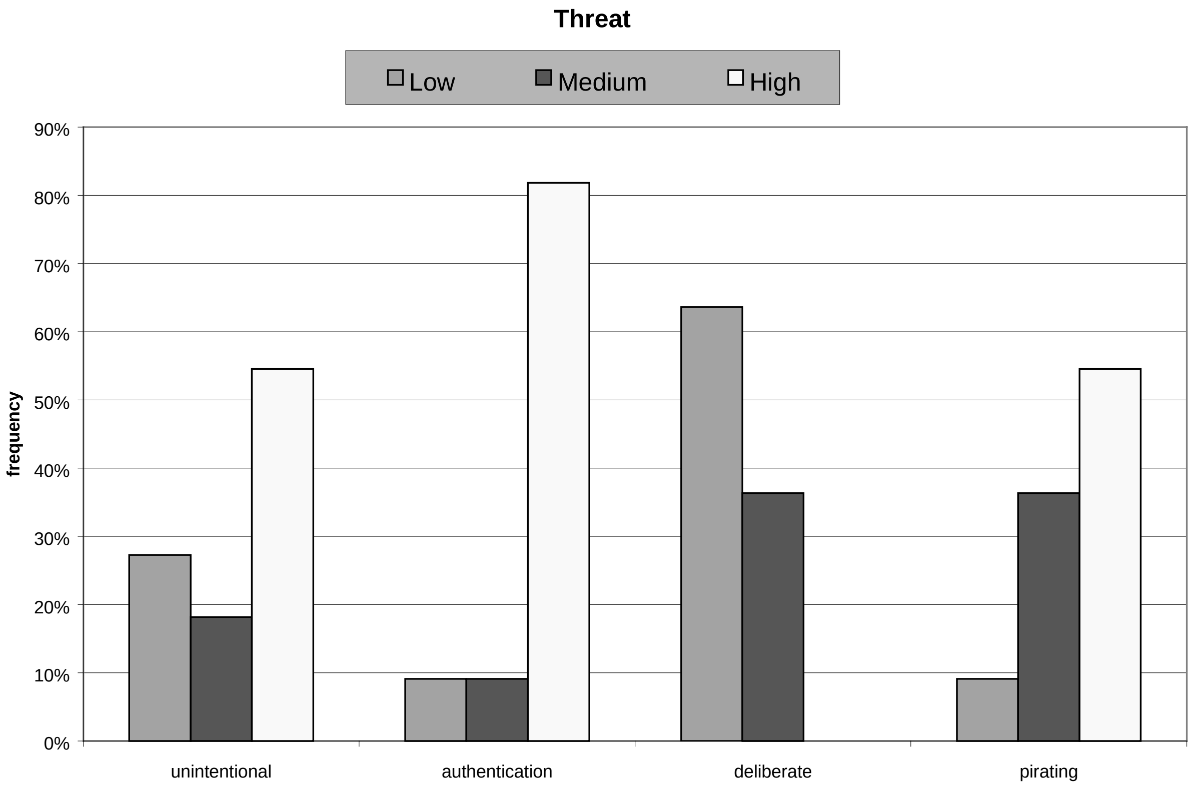
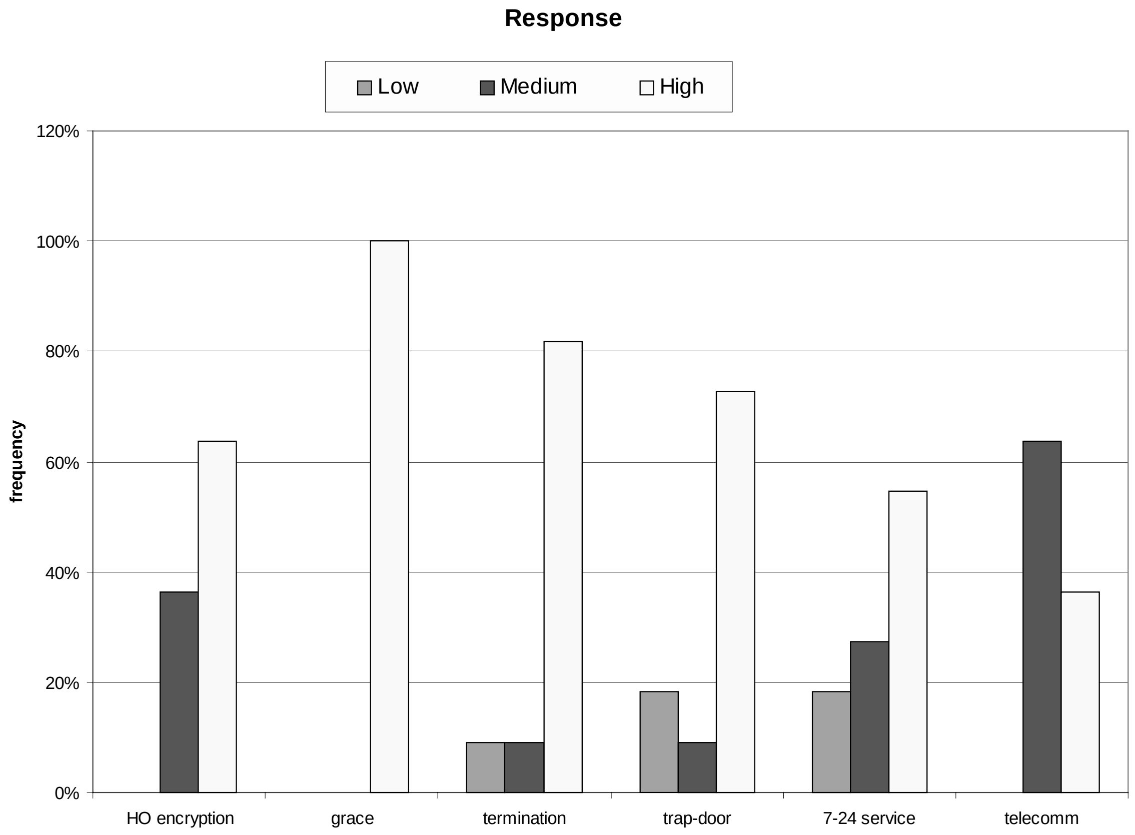


Figure 3 shows that HOs are likely to implement an encryption/authentication scheme to combat the perceived threats. A "grace period" will allow users access to the data for a period beyond the license due date but after that, access will be terminated. Some form of a yet to be determined "trap door" will allow users to access part or all of the data. A 7 days a week/24 hours per day service will likely be available to handle renewing license keys and the users are most likely be able to do this via some telecomm link at sea.

Figure 3: The Response



7) Immediate Options Available For A Security System

Should a HO or RENC decide to implement a security system in the near future there are two main options: adopt an existing encryption approach and implement it or have the data distributed through Agents in the SENC form.

7.1) The PRIMAR Model

7.1.1) How it works

The core technology of the PRIMAR encryption scheme is the BLOWFISH algorithm. This is a well known approach and provides more than adequate security. It is an approach used in many commercial implementations.

The basic system has been described in the report by Kibby & White to the UK HO in March 1999. The full report is published on the web site at

http://www.openecdis.org

The implementation of the UK algorithm was undertaken by PRIMAR and successfully integrated into their data management and management information systems. Engineering kits were developed and provided to manufacturers who wished to implement the decryption process into their ECDIS.

7.1.2) How it can be implemented

PRIMAR has provided some information of the basic structure of the encryption and authentication implementation. The security modules do not exist as a stand alone system and cannot be implemented in a “plug-and-play” mode but will require careful integration into the operations of each RENC.

PRIMAR will need to be contacted directly for details and support.

7.1.3) Notable Advantages

The security system has been designed and implemented by a leader in the industry and other RENCs can benefit from their knowledge. The security system itself provides state-of-the-art encryption and authentication technology.

7.1.4) Notable Cautions

Key Management requires a substantial effort. Implementing a foolproof method for keeping the system simple for the clients will require extensive planning and testing prior to a full roll-out of the security system. Since the PRIMAR model is not “plug-and-play” other RENCs will have to budget resources carefully for this task. The cost of maintaining the security system will have an impact on pricing.

7.2) SENC Distribution

7.2.1) How it works

In the SENC model the burden of security is placed on the supply chain. System manufacturers play a central role since the encryption and authentication functions are implemented at the SENC stage, not the ENC. Manufacturers and agents are free to choose whatever form of security system they wish, providing it meets some predetermined performance specifications which HO’s and RENCs can set.

7.2.2) How it can be implemented

The implementation details are left to the manufacturers and/or agents.

7.2.3) Notable Advantages

The HO’s and RENCs are spared the task of implementing a security system leaving more resources for base operations, ENC production and QA.

7.2.4) Notable Cautions

Not having implemented the security system the HO’s or RENCs are dependent upon the manufacturers and/or agents to provide the appropriate level of security. Clients are tied to one source for their data and thus do not benefit from commercial competition.

8) Overall Findings

Encryption is a complex issue with many conflicting requirements. It is not an endeavor to step into lightly.

PRIMAR's approach is state-of-the-art but does not achieve all security objectives. It is unlikely that any system would.

PRIMAR's model is portable but not "plug-and-play" and this might hamper its easy installation elsewhere.

Details regarding the transfer of the technology from PRIMAR to other RENCs or HO's remain to be determined.

Poll results show that even informed participants have widely varied opinions as to the benefits and implementation of encryption.

The SENC approach is viable, if only as a transitional one.

	Do Nothing	SENC	Watermark	Unregulated	PRIMAR	Change Standard
protection against corruption	N	H	N	H	H	H
demonstration of data integrity	N	H	H	H	H	H
demonstration of ownership	N	H	H	H	H	H
protection against corruption	N	H	N	H	H	H
protection against data piracy	N	H	N	H	H	H
Implementation simplicity	N	H	H	L	M	L
End-user simplicity	N	H	N	L	M	M
Speed of implementation	N	H	H	L	M	L

N = nonexistent
L = low
M = medium
H = high

Table 1: Table of Options

9) Additional Considerations For The Future

Encryption is not an Information Technological (IT) issue but a human factors one. The central problem remains the unintended effects of an overly aggressive solution.

9.1) Implement Partial Encryption:

Only encrypt part of the ENC allowing clear access to sufficient information to allow safe navigation. For example encrypt everything but the navigation layer. This involves changing the S57 standard.

9.2) Provide Low Level Encryption:

Although the encryption code is easily broken it is sufficient to act as a deterrent for most clients and provides sufficient security to allow legal action to be taken against those who pirate the data. Only time will tell if a strong security system is needed. If it isn't then a system easier and cheaper to manage can be implemented.

9.3) Simplify the Problem

One or more chart agents will likely implement automatic license renewal schemes. This might demonstrate that a strong security system is not required.

Reporting non-compliant ships to Port States might add some teeth to the license renewal problem.

9.4) Change the Problem:

Provide a Pay-Per-Use system wherein a full suite of ENC are provided unencrypted and users pay for each use. How to do the accounting for this should be possible with the data captured through the voyage tracking mechanism. An inventory of chart usage could be downloaded to the chart agent or RENC whenever updates are acquired through the Internet.

10) Recommendations:

10.1) Learn from Current Approaches

There are three security models in operation at the moment:

- No Explicit Security
- PRIMAR
- SENC

We have a unique opportunity to watch and see what the best approach is. Are the threats real ? Is there a measurable burden on the clients ? Do they care ? Are there serious leaks in the system ?

10.2) Keep Future Options Open

Other options, not feasible at the moment offer some potential for a modified security system. We need to keep these options open and discussible.

10.3) Re-Evaluate in Two Years

It is likely that within two years the situation will undergo some significant change. We will have to react to these changes in technology, market behavior or global politics.

Appendix 1: Primer on Encryption

Introduction and Some Terminology

Cryptography is the Science and Art of encoding secret messages, so that they cam be sent without being read by unauthorized third parties.   It is an attempt to come up with a *technological* solution to one of the problems of security.

This text will attempt to provide a brief explanation of what Cryptography is all about. It will not attempt to explain any ideas fully; nor will it attempt to provide working examples, or complete solutions. Its aim is to give a short glance, at what the field is all about.

Alice wants to send Bob a secret message. She could put the message in a steel Box, lock it shut, and hire armed guards to secure it on it's way to Bob, but instead she chooses to alter the message in some way, so that anyone unauthorized (Mallory for example) who does not know how she altered the message, will have a great deal of trouble reading (if it's possible at all) the message.

Though this may be obvious, it is worth noting that Alice and Bob had to have agreed on some protocol. That is Bob must know what Alice has done to the message, so that he (and hopefully only he) can make sense of it.

What Alice and Bob need is called Cryptography. What Mallory needs is called Cryptoanalysis.

Alice's message before she altered it is called the "Plaintext", and is often denoted by the letter P.   Her altered, hopefully secured text, that which Bob will receive, is called the "Cyphertext", and is often denoted by the letter C. The process of altering is called "Encryption", and is often represented as the function E(). The process of "unscrambling" is called "Decryption", and is often represented by the function D().

So E(P)=C, D(C)=P, so P=D(E(P))

Substitution ciphers

Perhaps our first suggestion to Alice should be that she should somehow "scramble" her message in a systematic way, so that then, only Bob will be able to unscramble it. Let's look at how this could be done. One simple way would be to shift all the characters by, say, 4. So that an "a" becomes an "e", a "b" becomes an "f", a "c", a "g", and so on. If she wanted to be a bit more creative she could replace characters arbitrarily using a table of replacements like a>c, b>x, c>n and so on. (She'd have to make sure that no two characters mapped onto the same letter, since then Bob wouldn't be able to translate back, but this is no big problem)

Bob to get the plaintext P, from Alice's cyphertext, C would only have to run the substitutions Alice made in reverse.

But Mallory isn't going to have to have many problems with this. The problem is that an English (and any other "natural" language text has just too many patterns that she can exploit to see how the message was "scrambled". And once she's figured out how one message (or a part of a message) was scrambled, then she can decode the rest of the message and all future messages, until Alice and Bob change their system.

For example, "a"s appear often alone or at the start of two letter words. ("a cat", "an apple").   "e"s appear often at the end of words. "t"s and "h"s often appear together, as do "c"s and "h"s.

Another strategy Mallory can use is to count the letter frequencies and then mach them up with the known letter frequencies of English (or whatever foreign language Alice and Bob talk). This isn't a perfect strategy of course, but Mallory will certainly gain many of letters this way.

And the more Alice and Bob rely on their scheme, the easier it is for Mallory to break it, since she'll have more text to analyze that way. If Alice writes Bob more than a few sentences this way, we shouldn't be surprised if Mallory breaks her system completely.

Real Cryptography

The above example shows how naive Cryptography works, and how cryptoanalysis is used to break cryptographic techniques. Cryptography works by "scrambling" a message so that (hopefully!) only authorized people can decrypt the message. Cryptoanalysis works by looking for patterns in the cyphertext, and using those patterns to figure out what the plaintext was.

Good cryptography makes the cyphertext as random as possible, so that no patters remain, that can be exploited by Mallorys.

Good cryptography algorithms are also hard to invent. There are no perfect solutions, but there are some good ones. Some of them are so good that major governments have tried and are trying to restrict their use.

Key Based Algorithms

One of the problems that Alice and Bob ran into when they decided on a cryptographic solution to their problem was that they had to invent a cryptographic algorithm. Not being Cryptographers, they were naturally afraid that if they chose a well know algorithm, then someone could simply guess which algorithm they chose and thus would be able to decrypt their messages. This concern would have been justified only in the case of non-key based algorithms.

The problem with non-key based algorithms, like the one Alice and Bob used is that the decryption, if not both the encryption and the decryption, algorithms must be kept secret. Since these algorithms cannot be shared, this means that for every cryptographic application, a different algorithm would be required. Good cryptography algorithms are hard to come up with, so when we find one, it would be nice if we could share the algorithm, without compromising, secrets, which we have protected, with the algorithm.

This is what we need Keyed Cryptographic algorithms for.

Suppose Alice and Bob decided to use a Keyed Cryptographic Algorithm, with encryption function E, and decryption function D. First they would generate a pair of keys KE, KD (which may actually be the same key depending on the particular encryption/decryption function they have decided on). Alice would keep KE and bob would keep KD. The algorithm could be public, but these keys would now be secret. These secret keys form parameters for the algorithms, which will assure that only Alice and Bob can read the messages that are encoded.

Now when Alice want to send message, P, to Bob, she computes C=E(KE,P), and sends C to Bob. Bob receives C and computes P=D(KD,C).

As long as the keys KE and KD are safeguarded the system remains secure, independent of whether the encryption functions E and the decryption functions D are public or become public.

This allows many people to use the same algorithms E and D, without being able to read each other’s secrets.

Private Key vs. Public Key Algorithms

Using a keyed algorithm allows Alice and Bob to use an encryption algorithm, without having had to invent it them selves.   But they still have a problem: they need to meet at some secure location to create or exchange the key. This doesn't seem like a big problem, until one realizes that if they could do that, then Alice could simply give bob the secret message instead them exchanging keys. Admittedly a set of keys allows many sets of messages to be exchanged, but still, it does seem, not only awfully inconvenient to arrange to meet in order to generate or exchange keys, it might be sometimes is impossible.

This is the problem that is solved by public key algorithms.

Bob wants to receive a secrete message from Alice, and Alice wants to send a secret message to Bob. They have decided they can't or don't want to meet at a secure place, to decide on what algorithm to use or exchange keys, so all their communication has to be done over an unsecured channel.

It seems odd, that in the end Alice will be able to send a secure message to Bob, but strangely enough it can be done.

Bob chooses a Public Key (or Public Key/Private Key, as they are sometimes called), Encryption/Decryption algorithm E()/D() (for example RSA or PGP), and generates a PUPLIC Encryption key, KE and a PRIVATE (!) Decryption Key, KD.

He tells Alice, over the unsecured channel which algorithm E()/D() he has chosen and what his PUBLIC encryption key, KE is.

Alice wants to send Bob her message, P. She computes the encrypted message, C, using E() and Bob's PUBLIC encryption key KE like this: C=E(KE, P), and sends C to Bob. Mallory, who has been eavesdropping on the unsecured channel, can't compute P. And this is true even though she knows exactly how Alice computed C using E() and KE and even D() doesn't help, since she doesn't know what Bob's PRIVATE KD is.

Bob receives C, and computes P=D(KD, C).

And Mallory is thwarted.

One Way Functions and Tractability

The odd thing about all this is that knowing E(), KE, and C does not help Mallory to discover P. That is, knowing everything about how the encryption was done, plus seeing the encrypted text doesn't help Mallory figure out how to decrypt the message.

This certainly seems counter-intuitive. But is it really?

You might be able to take a mechanical Swiss watch apart (Encryption) but without really good instruction (D() and KD) could you put it back together again? The real world is full of what cryptographers call one-way functions: Functions which are easy to compute, but who's inverse is hard to compute.

A common example is prime factorization: Consider the primes 3, 7, 7, 17, 19, 19, 19, 23, 29, 31, 33

It's pretty easy to multiply them all together; the answer is 1670823160083.

Now try to prime factorize it, the product.

You would need to write a prime factorization program and it will take considerable more computing power to factorize the product into its constituent primes. If the list of primes were a few pages long, you could still multiply them together in no time at all, but now all the computing power in the world would take more time than the universe has left to prime factorize it.

How one-way functions like this (and there are lots more) can be used to construct secure Public Key Encryption Algorithms is beyond the scope of this text. But it can be done, and there is no shortage of books that describe how.

A Few Words on Authentication

Encryption technology is all about making sure that an intruder cannot read an encrypted message.

Authentication technology is all about making sure that an intruder cannot alter or forge a message.

These two subjects, which are actually quite different, are often discussed together since they are both built on common technology. Any encryption algorithm can function as an authentication algorithm, in the following manner:

Alice wants to send Bob a message but Bob wants to make sure that the message he gets from Alice is really from her. Beforehand they have settled on Public-Key Encryption for their encryption standard so they each have a Private and a Public key.

Bob sends some secret Token, X, to Alice, which he encrypts with Alice's Public-Key. Only Alice can read this, since, of course, because only she can decrypt it using her Private-Key. She decrypts X, and re-encrypts it along with P, the message that she wants to send to Bob, but this time with Bob's Public Key. She now sends the whole package to Bob.

Now Bob decrypts the whole package using his Private-Key, but before he accepts the message's authenticity he checks to see if he the original token he sent to Alice is in the message. If it is, he knows the package must have come from Alice, since only she could have decrypted X (remember, he sent it out encrypted with Alice's Public key!)

A Good Book on Cryptography

This paper gave only a glimpse into this interesting subject. A great book, that can serve well both as a tutorial and as a reference on this subject is "Applied Cryptography" by Bruce Schneier, Published by John Wiley & Sons Inc. It is currently in it's second edition.