The Open Navigation Surface Project

This article is an updated version of a paper presented at US Hydro 2005 and reproduced with kind permission of the conference organisers.

Brian Calder1, Shannon Byrne2, Bill Lamey3, Richard T. Brennan4,5, James D. Case1, David Fabre6, Barry Gallagher1, R. Wade Ladner5, Friedhelm Moggert6, Mark Paton7 and the Open Navigation Surface Working Group

Abstract

Many hydrographic and oceanographic agencies have moved or are moving towards gridded bathymetric products. However, there is no accepted format to allow these grids to be exchanged while maintaining data and metadata integrity. This paper describes the Open Navigation Surface (ONS) Project, which aims to fill this gap. The ONS Project is an open-source software project designed to provide a freely available, portable source-code library to encapsulate gridded bathymetric surfaces with associated uncertainty values. The data file format is called a Bathymetric Attributed Grid (BAG). The BAG is developed and maintained by the ONS Working Group (ONSWG), and the source code is available via the ONS website.

Résumé

De nombreuses agences hydrographiques et océanographiques se sont orientées ou s’orientent actuellement vers les produits bathymétriques quadrillés. Toutefois, il n’existe pas de format accepté qui permette à ces quadrillages d’être échangés tout en conservant l’intégrité des données et des métadonnées. Le présent article décrit le projet ONS (Open Navigation Surface) qui vise à combler cette lacune. Le projet ONS est un projet de logiciel libre conçu pour fournir une bibliothèque portable de code source à libre disponibilité devant encapsuler des surfaces bathymétriques maillées avec des valeurs d’incertitude associées. Le format du fichier des données est appelé Carroyage bathymétrique attribué (BAG). Le BAG est développé et tenu à jour par le groupe de travail ONS et le code source est disponible sur le site Web ONS.

Resumen

Muchas agencies hidrográficas y oceanográficas se han orientado o se están orientando hacia los productos bathimétricos cuadriculados. Sin embargo, no existe un formato aceptado para que estas cuadículas sean intercambiadas manteniendo la integridad de los datos y los meta datos. Este artículo describe el Proyecto «Superficie de Navegación Abierta» (ONS) cuyo objeto es cubrir este vacío. El proyecto ONS es un proyecto de software de fuente abierta diseñado para proveer una biblioteca portátil de código fuente de libre disponibilidad para encapsular superficies bathimétricas cuadriculadas con sus valores de incertidumbre asociados. El formato de archivo de datos es llamado Cuadrícula Bathimétrica Tributada (BAG). El BAG es desarrollado y mantenido por el Grupo de Trabajo ONS (ONSWG), y el código fuente es disponible a través de sitio web de ONS.
Introduction

The Navigation Surface [Smith et al., 2002; Smith, 2003] paradigm is a design for a databased alternative to traditional methods of representing bathymetric data. It aims to preserve the highest level of detail in every bathymetric dataset and provide methods for their combination and manipulation to generate multiple products for both hydrographic and non-hydrographic purposes. The advantages of the method over traditional schemes (for example in optimisation of chart production, preservation of source data at highest available resolution, preserving a single database with multi-product creation capability, preservation and processing of uncertainty of data in addition to depths and integration of data from various sources) are such that a number of commercial vendors have adopted the technology. However, this multi-vendor situation means that there is a strong requirement for a method to communicate results in a vendor neutral technology. The Open Navigation Surface (ONS) project was designed to fill this gap by implementing a freely available source-code library to read and write all of the information required for a Navigation Surface.

The Navigation Surface concept requires that in addition to estimation of depth, an estimate of the uncertainty associated with the depth must be computed and preserved. In order to make the system suitable to support safety of navigation applications, we also require a means to over-ride any automatically constructed depth estimates with ‘Hydrographer Privilege’, essentially a means to specify directly the depth determined by a human observer as being the most significant in the area (irrespective of any statistical evidence to the contrary). Finally, we must provide data on the data, or metadata, which describes all aspects of the data’s life from methods of capture to processing methods, geospatial extents to responsible party. The ONS project provides means to incorporate all of these requirements in a portable, platform neutral, vendor neutral format.

The ONS project has as its primary goal to foster and support the development of a source level library to read and write the data format. A predicate of the project is that the implementation of the format in concrete terms through a generally available source library is more likely to engender rapid acceptance and adoption than a more formal approach (e.g., through a formal abstract standards proposal). By presenting the full source-code for the access library to public scrutiny, we hope to encourage contributions that benefit the project as a whole. The Open Source experience suggests that “Given enough eyeballs, all bugs are shallow” [Raymond, 2000], and we aim to gain from this mode of development.
This paper describes the design goals of the ONS project, the technology adopted to implement the goals, and the progress made towards a first release of the source code library. Details of the data encapsulation, metadata structure and digital signature schemes are also provided. Finally, we conclude with some perspectives on the future of the project.

Nomenclature: A Navigation Surface by Any Other Name

The term ‘Navigation Surface’ was coined to describe the combination of a data object representing the bathymetry and associated uncertainty of a localised area, and the methods by which such objects could be manipulated, combined and used for a number of tasks, including products in support of safety of navigation. These multiple goals have led to some confusion over what exactly constitutes a Navigation Surface. To avoid any further confusion, a revised nomenclature has been designed.

In the ONS model, a unit of bathymetry is termed a Bathymetric Attributed Grid (BAG). A single BAG object represents one contiguous area of the skin of the Earth at a single resolution, but can represent data at any stage of the process from raw grid to final product. The name Navigation Surface (NS) is reserved for a final product BAG destined specifically for safety of navigation purposes. The status of any particular BAG is distinguished solely by the certification section of metadata embedded in the file.

Basic Structure of a BAG

The basic aims of the ONS project are to generate a freely available, platform independent source code library for the data object, which supports large files and a hierarchical structure for the data. In order to provide these services, an abstraction and encapsulation layer is required. To avoid reinvention, the ONSWG adopted the Hierarchical Data Format [NCSA, 2005], HDF-V, which meets all of the core requirements, and is also multi-platform, supported, available in source code and binary form, and documented. The HDF-V layer also supports lossless compression of data if required, and has an extant base of tools to view and manipulate files, which helps to support development.

The mandatory components of a BAG are shown in Figure 1; the hierarchical structure of a BAG is shown schematically in Figure 2. The HDF-V implementation follows this pattern directly, except that the ‘Certification’ section is not included. To be considered valid, a BAG must contain a metadata element, a bathymetric grid, an uncertainty grid and a change-list indicating any modifications made to the grid due to hydrographic concerns.

![Figure 2: Logical structure of the BAG file format and digital signature certification segment.](image)

Layer-specific meta-data is limited in nature, for example numerical bounds on the data; file-specific meta-data should contain all information required to represent the file, including capture parameters, processing specification and certification of the data.
Each of the elements (except the metadata) can have element specific metadata (e.g., minimum and maximum elevation in the grid). The meaning of 'uncertainty' has not been defined directly in the base definition of a BAG model, since it may be different for different users, and may change through the lifetime of the BAG from raw data to final product. (For example, the uncertainty of a base BAG contains information about the raw data processed, and is intended for use by hydrographers and cartographers in making compilations; uncertainty in a product BAG is intended to indicate reliability of the depths for the end user, and may be significantly different.)

The 'Hydrographer's Modifications' change-list element can be empty. Otherwise, it contains the original depth and uncertainty values for every node that has been modified in the BAG due to operator intervention. This is a required feature to ensure hydrographer override and hence safety of navigation. To ensure that the base BAG product is safe, the values in the elevation grid are the modified (safe) values; since these are read unless the user specifically requests the original value, the BAG is fail-safe.

The BAG's internal structure allows for extension components to be defined at the same level of the file as the mandatory elements. Extensions will be considered as they are proposed by users – for example, to allow an alternative elevation layer corrected assuming a fixed speed of sound (this is common with data to be used with simple fathometers). The only restriction is that the full structure of the proposed extension must be provided to the ONS Architecture Review Board (q.v.) before adoption into the format definition.

In order to avoid confusion over interpretation, a number of axiomatic definitions are used in a BAG structure, Table 1. The definitions are essentially arbitrary, but avoid over-flexibility in the data structures, helping to keep the BAG a simple format to read and write. In the spirit of simplicity, projected grids are used by default, and are arranged in row-major order with the first element in the file being the southwest corner of the grid (i.e., the data progress west to east, and south to north within the area). Each node in the grid is taken to represent a point estimate of the respective data element at the appropriate location. Hence, the georeferencing for the grid consists of a single value indicating the position of the southwest-most node in the grid. A right-handed coordinate system is used throughout, resulting in increasing vertical values indicating shoaler depths, and negative values for all locations below the vertical datum. To avoid confusion, the term 'elevation' is preferred to 'depth' in this context.

<table>
<thead>
<tr>
<th>Element</th>
<th>Units</th>
<th>Reference</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement Units</td>
<td>SI</td>
<td></td>
<td>Only metric SI units are used in BAGs</td>
</tr>
<tr>
<td>Horizontal Datum</td>
<td>-</td>
<td>NAD83, WGS72, WGS84</td>
<td>Currently, all vertical datums are chart datums. A datum based on ellipsoidal heights might be added in a future version of the library</td>
</tr>
<tr>
<td>Vertical Datum</td>
<td>Metres</td>
<td>ALAT, ESLW, HAT, HTL, ISLW, LAT, LLW, LNWL, LWD, MHHW, MHLW, MHW, MHWN, MHWS, MLHW, MLLOW, MLLWS, MLW, MLWN</td>
<td></td>
</tr>
<tr>
<td>Projections</td>
<td>Metres</td>
<td>Mercator</td>
<td>General use worldwide</td>
</tr>
<tr>
<td></td>
<td>Metres</td>
<td>TM/UTM</td>
<td>General use low/mid latitudes</td>
</tr>
<tr>
<td></td>
<td>Metres</td>
<td>Polar Stereographic</td>
<td>General use high latitudes (over ±64°)</td>
</tr>
<tr>
<td></td>
<td>Metres</td>
<td>Lambert Conformal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Degrees</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Seconds</td>
<td>1970-01-01/00:00:00 UTC</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations:
Vertical Datums: A: Astronomical; D: Datum; E: Equatorial; H: Highest or Higher; I: Indian; L: Lowest or Lower; M: Mean; N: Neap; S: Spring; W: Water. Exceptions: ALAT: Approximate LAT; HTL: Half Tide Level; LNWL: Lowest Normal Low Water.
Projections: TM: Transverse Mercator; UTM: Universal TM.

Table 1: Axiomatic definitions for the bag data object.
Only a limited number of horizontal datums and projections are defined, although the vertical datum selection is more flexible since this choice does not require any code to be generated in the library. Use of unprojected grids is allowed, and limited transformation between geographic and projected coordinate spaces is included in the sense that point queries into the data can be made in either mode irrespective of the native coordinate system in the BAG. Reprojection of the grid (for example from UTM to Mercator, or UTM to unprojected) is, however, outside the scope of the base library. Irrespective of projection mode, all of the parameters required to fully define the coordinate system are required to be stored in the metadata. Projections are carried out using the US National Geospatial-Intelligence Agency’s GeoTrans package [NGA, 2005]. The list of datums in Table 1 is not intended to be finally exhaustive, and may be extended for other regions of the world as required. Mechanisms for this extension are part of the ONS Architecture Board, which is defined later in this paper.

**Metadata**

Metadata is essential to the value of the data, and is a mandatory part of the BAG format; the BAG format defines metadata as all auxiliary information required to interpret the basic data itself. This includes not only the details of the data collection (the who, what, where, why and when), but also the processing stages applied to the data since capture and legal statements about the validity, intended use, or expected disposition of the data in the BAG. (These certifications are linked to the security and authenticity mechanisms that protect the integrity of the BAG, which are described in the following section.) The metadata is embedded in the BAG file itself, rather than being a separate file. This is intended to ensure that the metadata and the data cannot be separated.

Models for geo-spatial metadata are now common. The BAG requires the use of a profile of the ISO 19115 standard [ISO, 2003] for the content of the metadata, Figure 3, and the proposed ISO 19139.
standard [ISO, 2005] for implementation in XML [Bray et al., 2004]. XML allows the metadata to be stored and retrieved in a system independent format, and provides a machine and human readable version of the metadata at no extra cost. The XML can be viewed and edited in a simple text editor (see, e.g., Figure 4), or more sophisticated displays can be created by using any XML data parser, many of which are freely available. The master copy of the BAG XML schema can be retrieved from the ONS website [Lamey et al., 2004], and is used internally to validate BAG metadata as it is read or written. To keep the BAG API simple and flexible, and to avoid exposing any third party XML parsing library, the interface for retrieving or providing the XML metadata uses simple UTF-8 character streams [Yergeau, 2003]. A cross-walk between ISO 19115 and the US Federal Geospatial Data Committee (FGDC) metadata standard is available (Pearsall, 2001), but is not supported by the BAG specification.

Veracity, Integrity and Authenticity

Digital data products have very different characteristics to traditional paper products. In particular, since there is no physical context, traditional methods of authorising editions and preserving the integrity of data are not applicable. Unless a breach of physical security takes place, the longevity and integrity of a mylar smoothsheet (or fairsheet) are almost guaranteed. The same cannot be said for digital archives. Gridded bathymetric data is very dense in the sense that the ratio of raw data to data structuring elements is very high. Therefore, whereas a randomly placed single byte error would probably be detected in vector data (e.g., S-57 ENCs) due to the fact that it would most likely occur within a structuring element and the data would fail to load, a similar error in a large gridded dataset would most likely occur within the data itself and hence be undetectable. Significant (e.g., more than 20 years) longevity of digital data

```xml
<?xml version="1.0" encoding="UTF-8"?>
<MD_Metadata xmlns=http://metadata.dgiwg.org/smXML
xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
xsi:schemaLocation="http://metadata.dgiwg.org/smXML
C:\ISO\smXML\metadataEntity.xsd">...
<extent>
  <EX_Extent>
    <geographicElement>
      <EX_GeographicBoundingBox>
        <extentTypeCode>true</extentTypeCode>
        <westBoundLongitude>-75.0</westBoundLongitude>
        <eastBoundLongitude>-74.0</eastBoundLongitude>
        <southBoundLatitude>46.5</southBoundLatitude>
        <northBoundLatitude>47.5</northBoundLatitude>
      </EX_GeographicBoundingBox>
    </geographicElement>
  </EX_Extent>
</extent>
<certification>
  <CI_ResponsibleParty>
    <individualName>CAPT. S. B. Collins</individualName>
    <role>ChiefOfDivision</role>
    <keySource>http://chartmaker.noaa.gov/certificates/index.html?keySource>
  </CI_ResponsibleParty>
  <CT_Certification>
    <certification>This data is the legal product of the survey</certification>
    <certification>This data may be used for product creation</certification>
  </CT_Certification>
  <CT_SequencelD>
    <seqid>1</seqid>
  </CT_SequencelD>
</certification>
<dateStamp>2001-11-28</dateStamp>
</MD_Metadata>
```

Figure 4: Example XML encoded meta-data extract. The meta-data is human-readable and editable in any text-editor (although specialist XML editors also exist), and contains information on data capture, processing, spatial extents, intended use, responsible party, etc. The XML string is simply encapsulated in the HDF-V structure.
is still essentially an unsolved problem, not because the data itself cannot be copied perfectly, but because typical archival physical storage media either decay or become unreadable due to obsolescence. Technical solutions for longevity will emerge almost surely (e.g., systems that automatically migrate data from obsolete disc-based spinning store to new discs as hardware is added), and therefore we consider here the problems of showing veracity, integrity and authenticity of BAG data objects.

Since the final intent for BAG objects is as legal archives for hydrographic data and products to support safety of navigation, it is essential that there is a way to prove that a BAG has been authorised by someone with appropriate credentials, and that it has not been modified (either intentionally or through transmission errors) since it was so certified. These are common problems in electronic commercial transactions, for which standard solutions are readily available. Based on algorithms for public-key encryption [Schneier, 1996], the US Federal Information Processing Standard (FIPS) Digital Signature Standard (DSS), FIPS 186-2 [NIST, 2000] computes a number from the file to be signed using the Secure Hash Standard (SHS), FIPS 180-2 [NIST, 2002] and combines it with some information known only to the signer (the private or secret key) in order to make the signature (a very large number). The process is shown schematically in Figure 5. This signature has the property that it is very difficult (essentially impossible) to compute without knowledge of the secret key. However, if you have the complement of the signer’s secret key (known as the public key), then it is easy to verify that the person who affixed the signature to the file did in fact have knowledge of the secret key, Figure 6. The signature and file are inextricably linked: if the signature is modified, it will not verify; if the file is modified, it will not verify. Hence, the DSS also provides protection from errors due to transmission and ensures that a grid prepared for safety of navigation purposes cannot be falsified without detection. There is no standard implementation of either DSS or SHS, and the US National Institute of Standards and Technology (NIST) will only certify compiled executables, and not source code. To implement the encryption primitives required for the algorithm and maintain cross-platform source code support, the BeeCrypt library [Deblier, 2005] was used. Note that although the DSS is based on cryptographic principles, the data itself is not encrypted.

The BAG specification limits the certification to the whole of the file, rather than attempting to protect individual components of the file; the certification itself only assures the file, leaving the implication of the certification to the contents of the metadata. It is a requirement of the format that before the BAG is digitally signed to certify the contents, an element

---

**Figure 5: Principles of digitally signing a BAG.** The raw data is passed through a filter to form the File Digest, which is then combined with some Secret number known only to the signer in order to form the Signature to be appended to the raw data. The probability of two files having the same File Digest is vanishingly small, and hence any modification to the file will cause the File Digest and therefore signature to change.

The Signature is also characteristic of the signer’s Secret number, and therefore can only be affixed by the signer.
Table 2: Contact location and administration points for the ONS project.

<table>
<thead>
<tr>
<th>Element</th>
<th>Location</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Website</td>
<td><a href="http://www.opennavsurf.org">www.opennavsurf.org</a></td>
<td>Description of the project; source snapshots; documentation; news</td>
</tr>
<tr>
<td>CVS server</td>
<td>cvs.ccom.unh.edu</td>
<td>Source repository for library</td>
</tr>
<tr>
<td>General Mail</td>
<td><a href="mailto:navsurf_general@ccom.unh.edu">navsurf_general@ccom.unh.edu</a></td>
<td>General discussion of ONS; Policy and Administrative details</td>
</tr>
<tr>
<td>Development Mail</td>
<td><a href="mailto:navsurf_dev@ccom.unh.edu">navsurf_dev@ccom.unh.edu</a></td>
<td>Developer discussion and co-ordination; includes <a href="mailto:navsurf_general@ccom.unh.edu">navsurf_general@ccom.unh.edu</a> traffic</td>
</tr>
<tr>
<td>Join Requests</td>
<td><a href="mailto:navsurfjoin@ccom.unh.edu">navsurfjoin@ccom.unh.edu</a></td>
<td>Requests to join either mailing list</td>
</tr>
<tr>
<td>Source Requests</td>
<td><a href="mailto:navsurf_access@ccom.unh.edu">navsurf_access@ccom.unh.edu</a></td>
<td>Requests for access to the CVS repository</td>
</tr>
</tbody>
</table>

is added to the metadata indicating the intent of the signature. This might be as simple as ‘data is complete and verified’ to ‘suitable for compilation’, ‘suitable for database insertion’, or even ‘certified for navigational use’. Presence of a signature and certification in a BAG does not imply that the signing entity has the appropriate level of authority to make the attached certification. Verification of a signing entity’s credentials is auxiliary to the BAG structure, and is not considered. Implementation of a certificate scheme, key management and other details of a public-key infrastructure required to support a digital signature scheme [Schneier, 1996] are also not defined in the BAG structure. These are considered a matter for the adopting agency, but a number of standard schemes are available. One possible scheme is provided with the ONS library, which uses a standard USB hardware dongle [Aladdin, 2003] to hold the secret key, encrypted using the Advanced Encryption Standard (AES) algorithm, FIPS 197 [NIST, 2001], and presents the public key as an XML-based certificate so that it can be embedded in the metadata if required [Calder et al., 2005].

Support and Administration

The BAG format is defined through its format document [Moggert et al., 2004], and supported through a web-site (URI: www.opennavsurf.org). The source code for the BAG library is developed collaboratively using a number of e-mail lists for coordination (Table 2), and is generally available from a Concurrent Versioning System (CVS) server to ensure consistency between the developers. Source code snapshots for users who prefer the last known stable release of the library are available from the web-site. The library is implemented

Figure 6: Verification of a BAG’s digital signature. The File Digest is computed as before, and then combined with the stored signature and the purported signer’s Public number (which is the counterpart of the Secret number used in signature). Easily checked mathematical relationships between the three numbers confirm whether (a) the purported signatory did in fact know the Secret number at signature time, and (b) that the file and/or signature have not been modified in the meantime. Interpretation of the consequences of this implication is a matter for the user application, based on the contents of the BAG’s metadata.
using the ANSI C standard, with primary target platforms of Linux and Win32 with secondary targets of other Unix systems as time and interested volunteers are available.

The ONS is a community-led effort, and the BAG format is supported by the volunteer support of the ONSWG. The format is maintained through an Architecture Board (ONSAB) co-opted from the members of the ONSWG and potentially others as time progresses. The ONSAB coordinates bug reports, change requests and feature improvements in the BAG format, and accepts requests for and definitions of new sub-objects from the public. The functions of the ONSAB are defined through the BAG Format Specification Document [Moggert et al., 2004].

Summary

Advances in all-digital hydrography with databased products have potential for significant advances in productivity, but this potential is unlikely to be realised unless vendors and users share a common vocabulary to describe the data objects being created and manipulated. This language must also recognise and express the differences between a digital form of a physical product, and a digital product with no physical context.

The Open Navigation Surface Project was established to design and develop a specification for hydrographic gridded data, and to implement a freely available source code library to support the specification. The essence of the BAG effort is to develop a solution that can be used now and grown organically. Restriction in complexity, such as those embodied in the axiomatic definitions of Table 1 for example, are a manifestation of this goal. This does not preclude standardisation at some point in the future.

An ONS object is known as a Bathymetric Attributed Grid (BAG), and consists of metadata, a grid of elevation data, a corresponding grid of uncertainties describing the elevation data, and a list of modifications made to the grid by a hydrographer in support of safety of navigation. Extra information may also be present in the file if required. After the body of the data, a digital security system block is appended to provide verification and authentication services for the whole of the file. All of the components of the BAG are built using standard libraries, methods and technologies. All of the components can be built from source, and full source code is provided.

The project is supported by the community for the community; participation in the project (c.f., Table 2) is positively encouraged.

References


Bray, T., J. Paoli, C. M. Sperberg-McQueen, E. Maler, F. Yergeau (Eds.), 2004. Extensible Markup Language (XML), 1.0 (3rd Ed.). (www.w3.org/TR/REC-xml)


Raymond, E. A., 2000. The Cathedral and the Bazaar. (www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar [XHTML; note that other versions, including PostScript and XML are available from the same location.]).


Biography of the Author

Dr Brian Calder is a Research Assistant Professor at the Center for Coastal and Ocean Mapping & Joint Hydrographic Center at the University of New Hampshire. He graduated MEng and PhD in Electrical & Electronic Engineering from Heriot-Watt University in Edinburgh, Scotland in 1994 and 1997 respectively, but was subsequently seduced into sonar signal processing for reasons that are now obscure. His research interests have previously covered speech recognition, texture analysis, object detection in sidescan sonar, high-resolution sub-bottom profiling, simulation of forward-looking passive infrared images, acoustic modelling and pebble counting. Currently, they revolve around the application of statistical models to the problem of hydrographic data processing.

E-mail: brc@ccom.unh.edu