

AN ON-LINE SERVICE TO ACCESS AND VISUALIZE THE DATA ALIGNMENT LAYER

B. Ameri, A. Shankai, R. Orth

MacDonald Dettwiler and Associates (MDA)
13800 Commerce Parkway, Richmond, BC, Canada
Tel: (604) 231-2308, Fax: (604) 278-2117
bameri@mda.ca, ashankai@mda.ca, rorth@mda.ca

Commission IV

KEY WORDS: Web mapping, On-line services, Spatial Search, HTML/XML, RDBMS

ABSTRACT:

The CGDI (Canadian Geospatial Data Infrastructure) currently supports access to the CDAL (CGDI Data Alignment Layer) data via a public Web site, however, the process to retrieve the data is designed for human interaction and is not available programmatically as an online service. Furthermore, retrieval of individual, or an arbitrary range of points is not currently supported by the CDAL, and instead complete data sets have to be requested, where a data set consists of a regular area of points based primarily on the NTS (Canadian National Topographic Sheet) system. A data set can be retrieved only one at a time by specifying its data set name or number.

This paper describes a new service that supports more general access and visualization of the CDAL data. The new CDAL system is an online service and provides easy access to a database of ground control points. The clients of this service are able to obtain responses in either HTML or XML format and to request CDAL data in either a textual or graphical form. The service accepts HTTP requests that permit a geographic region of interest to be specified as the search criteria. Only those points found within the region of interest will be returned. The second service, built on the first one, uses the textual information associated with CDAL ground control points to deliver rasterized image chips rendered for visualizing points of the CDAL.

1. INTRODUCTION

This paper describes a new service that supports more general access and visualization of the CDAL data. The new CDAL service (CDAL 2001) is an online service and provides easy access to a database of ground control points originating from the 1:50,000/1:250,000 Canada national topographic base map series (NTDB data sets), which is compiled by the Center for Topographic Information (CTI)-Sherbrooke. All that the user needs to access CDAL service are an HTML enabled browser (e.g., IE or Netscape) and Internet connection.

The new CDAL service is an online bilingual service and provides easy access to a database of 7 million ground control points. The database stores textual items in both English and French. The contents of these two data sets are identical, with exact geometry but different attributes in terms of language.

Using the existing CDAL service provided by CTI-Sherbrooke (Parent 1999, CDAL 2000) the user can only retrieve data sets based on a data set number, which corresponds to different areas where the control points belong. This method of information retrieval is not efficient since there are cases where the user wishes to retrieve the data corresponding only to a small region of interest (or even a region larger than a typical area, for that matter).

Another drawback of the current service is that the user usually receives too many records and the main task of examining the data and extracting the more appropriate data items will be left to the user as a manual operation. The new CDAL service will

allow for specifying a region of interest as well as the type of control point (feature type) for information retrieval. The client will be able to obtain responses in either HTML or XML format. An image chip featuring each identified control point in the region of interest will be displayed in GIF format. Also, another image chip in GeoTIFF format will be available if the user wishes to fetch and store it on their local machine. Such a service is of economic benefit e.g., for the production of geo-coded satellite/aerial imagery, and is geared towards clients who need to use the CDAL for activities such as ground control point marking where ground control is often expensive and difficult to obtain.

An overview of the new CDAL service is presented in the next section. Section 3 discusses the system architecture. The technology used in this project, particularly existing COTS product, and a detailed discussion on the development of the system and data-flow are given in sections 4 and 5 respectively, followed by a conclusion in section 6.

2. OVERVIEW

CDAL is an online bilingual service and provides two new online services for the CGDI:

Spatial Search engine: The CDAL service allows the user to perform a spatial search and retrieve all, or only selected categories of the CDAL entities that belong to the specified geographical region. The search mechanism is capable of accepting the user-defined parameters, find the required element(s) based on the specified criteria, and pass the result

into the visualization module. Figure 1 shows the input area of the CDAL front page where the user can specify the search criteria.

Figure 1: CDAL front web page, where a query can be formed and submitted

The input area is divided into three main sections. The first section defines the Region Of Interest (ROI). The user specifies the bounding rectangle of the area for which a search should be conducted. The bounding rectangle is defined by entering the geographical co-ordinates of its upper left and lower right corner. The second section is for specifying the scale. The control point data stored in the database are derived from two different map scales (1:50000 and 1:250000). The user has the option of choosing either or both of the scales. And the third section is where the user selects the type of control point(s) to retrieve from the database. There are four types of control points to search. These are polygons, lines, intersections, and points. The most common example of a polygon control point is a lake. A bridge is an example of line control points and a well or a water tank can be considered a point. Intersections are comprised of 3 or 4 branches of roads, railways or highways (see 5.1 for more details).

Figure 2: Snapshot of CDAL front page containing the outputs

Visualization and Rendering: The CDAL service is able to receive, reformat and visualize the results of the search process in a textual (HTML/XML) and graphical (GIF/GeoTIFF image chips) format. Assuming that a query is submitted and a set of ground control points within the specified search area is found, then a rendering function will be triggered, which reads the textual information of the CDAL data elements and generates two contextually identical image chips. These image chips consists of a thumbnail GIF image for displaying on the front web page and a GeoTIFF image chip stored on the server ready for download by the user. Figure 2 shows an example of the textual information and the corresponding image chip displayed on the output area of the CDAL front page. At the beginning of the display page is a summary of the accumulated total number of records as well as the total number of records found for each selected feature type. The user can view the text information and the associated GIF image chip as the standard output. However, the user may perform extra actions for retrieving the GeoTIFF image chip, or the XML record of the result.

CDAL output can be used in combination with other systems to perform a specific task such as image geo-coding. It is important to note that CDAL does not interface automatically to any other system, and there will always be some level of human intervention to use CDAL output as an input to another system.

An immediate application of the service is the incorporation of the ground control points into an efficient and cost-effective geometric correction of Earth observation data. The textual descriptions provided by CDAL can be extremely helpful in identifying and marking control points. There is an overall increase in operator efficiency because control does not have to be digitized from paper map products, but rather retrieved from CDAL.

3. SYSTEM ARCHITECTURE

Figure 3 illustrates a context diagram for the CDAL system including online operation and conversion/update operation. It mostly focuses on the external interfaces of the system. During the online operation, the user communicates with the system through an HTTP link. Upon receipt of a request (query) from the user, the application runs the query against the CDAL database and returns the appropriate information to the user.

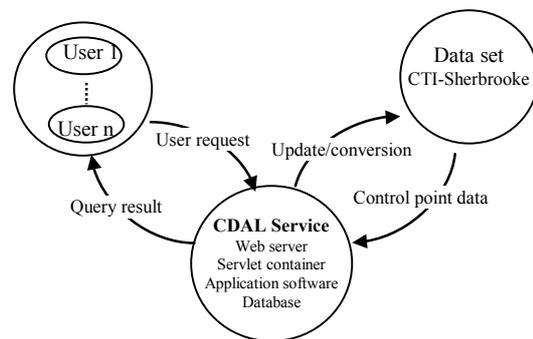


Figure 3: CDAL system context diagram

Conversion/update operation takes place whenever it is required to extract new information from the CGDI data set and insert it into CDAL database. It is not part of the online service and can be considered an administrative operation. The data stored on CGDI data set is in ASCII format. The conversion/update module extracts the control point data, converts it to suitable

format for CDAL database and inserts it into CDAL database. This mechanism is performed once at the beginning of the CDAL operation (conversion) and thereafter, in a regular basis (see section 5.1 for more details).

Figure 4 illustrates the on-line operations and the inter-relationships between CDAL components. Operations depicted by dashed line are optional operations that may or may not occur depending on the user request.

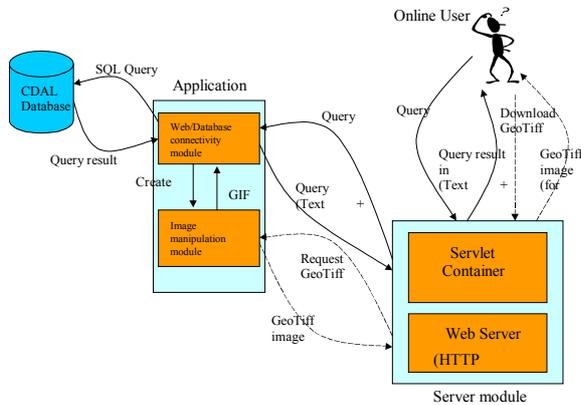


Figure 4: Operational overview of the CDAL service

The user requests the Web front page of the CDAL service by entering the CDAL URL into the web browser. Then s/he specifies a set of constraints to form a query and submits the query to the CDAL server. Upon receipt of the query, the application connects to the CDAL database to retrieve the data that satisfy the query. The result is prepared in Hypertext Markup Language (HTML) format and is sent back to the client.

Apart from the regular information that the user receives, s/he may also wish to receive additional information such as a GeoTIFF image featuring a specific control point (dashed lines). This image chip can be used for different application such as image geo-coding. By default, the information is provided in HTML format but the system allows the user to receive the information in eXtensible Markup Language (XML) format as well. The HTML format is useful if the response is to be displayed as a Web page. The XML format adds structure to the CDAL data making it easily accessible for further processing by client applications.

A detailed description of the above components and the technology used in this work is presented in the following section.

4. TECHNOLOGY LAYERS

The CDAL service consists of four main components that form the technological layers of the system: the Web server, the application server (servlet container), the database, and the application software (see Figure 5).

These are the minimum system components to fulfil all the requirements for a system such as CDAL. The web server and servlet container are used primarily for web-enabling the application. By web-enabling, we mean providing the service over the Internet so that a remote client can connect to the CDAL server, submit a query based on some pre-determined

constraints, and receive the appropriate information in the desired format.

The server side components of the system are a combination of commercial off-the-shelf (COTS) products, which provide the basic framework for developing the CDAL on-line service. The applications software was developed to implement the specific requirements of the system and is mainly responsible for receiving and processing the on-line requests, data extraction, content creation, and display. The applications software serves as the central component that interfaces to all other components. In addition a data conversion module was implemented to convert the original textual data sets into a format suitable for a RDBMS. The following is a brief description of the above components.

Web Server: this component hosts the Web pages, receives the incoming http requests and passes them to the Web application server (servlet container) to be processed. Our choice for Web server is Apache Web server, which can be downloaded and installed with no cost and has proved to be a reliable. Currently, close to 70% of Web servers around the world are using Apache (Apache 2001).

Application server (servlet container): This component is used to create dynamic Web pages where the flow of information is bilateral. That is, the user can send information to the Web server, which can be used to provide the content requested by the user. In CDAL context, the user specifies the criteria for a query, which will be received by the server and will be used to perform the appropriate action (Tomcat 2001)

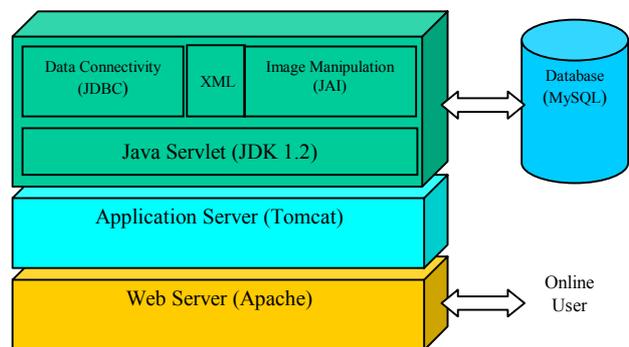


Figure 5: Technology layers of the CDAL service

Database: The original CDAL data points are organized based on the National Topographic Sheet (NTS) number in a flat file data structure. Each file consists of a set of points, where every point is categorized by a series of attributes. These points are not related to each other and are not organized into a DBMS. Also, accessing and selecting an individual point from CDAL data is not supported. To overcome this problem a relational DBMS is utilized to restructure the CDAL data into a suitable format where accessing individual points is possible. There are two series of tables in the CDAL database (see Figure 7 for a conceptual representation of GCP tables and their relationships). The First group of tables store the ground control point (GCP) data and are actually part of the CDAL core operation, whereas the other table maintains a history log that keeps track of the incoming queries, and is mainly used for bookkeeping and management purposes.

Application software: In addition to the above layers, a software layer must be implemented to deploy the CDAL specific tasks. The elements of this layer are image manipulation, XML support, database connectivity, and capturing the online queries. Additionally, an HMI (Human Machine Interface) module has been implemented as part of the CDAL-specific software. This involved content creation for the HTML response pages. A detailed description of the developed software modules are discussed in the next section.

5. SOFTWARE MODULES AND DATA FLOW

This section introduces the main software modules that comprise CDAL. Some of these modules are standalone whereas some are used in the context of a specific software component. For example the database connectivity module operates in the context of the servlet container. But the image rendering module can be seen as a stand alone module. The bookkeeping module can be executed completely independent from the operation of the CDAL on-line service. The following is a list of all software modules that are implemented in accordance with the software components.

- 1- HTML web front page
- 2- Image rendering
- 3- Database connectivity
- 4- Result web page (HTML (default) or XML)
- 5- Bookkeeping / statistics
- 6- File cleaning (to delete all the temporary files such as image chips, XML files, etc.)

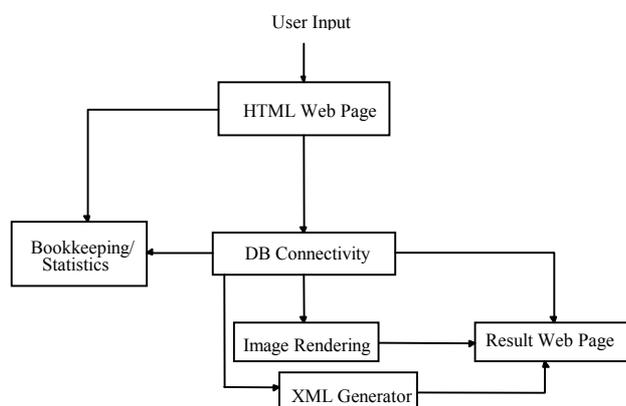


Figure 6: Software Module and data flow

Figure 6 illustrates the data flow and interrelation between the above modules. It shows the control flow between the modules in terms of input/output and not the timing sequence of the operations. The file cleaning program is not included in this figure because it does not interact directly with any of the system components. It runs in parallel with, but independently from, the CDAL server.

The *HTML web page* receives the user request and passes it to the appropriate servlet class for processing. In the CDAL system, the user's request is a number of constraints that comprise a query. This query is applied to the control point database, causing the database connectivity module to start operating. This is implemented using Java Database Connectivity (JDBC) functions, which are called from within a Java servlet class.

Another module that receives input from the front Web page is the *bookkeeping and statistics* module. The main responsibility of this module is to keep track of all hits to the Web page and the actions performed on the database, and to store the information for future reference. As the Web front page receives requests from users, it notifies the bookkeeping module of the requests so that it updates a counter and stores information (e.g., the IP address of the remote user).

The *database connectivity* module provides input for four modules, namely the bookkeeping module, the image rendering module, the XML generator, and the result Web page module. The database connectivity module can be considered the core of the CDAL system where all the useful information is extracted. This module notifies the bookkeeping module about the number of records returned for each request. It also provides the image-rendering module with information required for drawing the appropriate feature. The XML module uses the output from the Database Connectivity module to create the XML equivalent of the record. Finally, the output of the database connectivity module is used to create the text information to be displayed in the result Web page.

The main responsibility of the *image-rendering* module is to create image chips that graphically represent the control point data. Currently, two image chips with GIF and GeoTIFF formats are created. The GIF format is included in the result Web page whereas the GeoTIFF image is available for download. The module reads the geographical information such as geographical co-ordinates of the resulted point, which are stored in a text file as metadata, and then attaches them to a standard TIFF image chip, thus creating a GeoTIFF equivalent of the original GIF image.

The *XML generator* module is a static class that is responsible for creating an XML file for every control point that is contained in the result, using the specified Document Type Definition (DTD). There is a hyper link to the XML file for every control point in the result Web page. The user has the choice of downloading the XML file, or if the browser is XML enabled, the file could be viewed on the browser. The contents of XML tags are dependent on the language of choice but the tags themselves are not bilingual.

The *result web page* module creates an HTML page that can be viewed by the user's browser and includes textual information regarding the control points that have satisfied the user criteria. The user can proceed by requesting more specific information, namely the GeoTIFF representation of the control point and the XML equivalent of the HTML page.

5.1 Data Conversion

As mentioned before, this process is not part of the online service and can be executed in an off-line operation as part of administrative process. However, it is a major part of the CDAL application software and therefore it is discussed separately.

The original data set is currently stored in text format in two separate files, English and French language, and need to be converted into relational data in accordance with the proposed data model (see Figure 7). It is an important issue since the same implementation of conversion is used later for data updates when data entry is made into the textual data set. The data conversion is a two-step operation. The first step is inserting new records into the database by reading from the

English data set. This is seen as the record generation step. During this step, the French fields in the record will remain blank to be later filled in the second step. The second step is reading through the French data set, retrieving the French-specific fields, and updating the French fields in the corresponding record of the database. A unique data item is used to find the corresponding records in the French and English data sets.

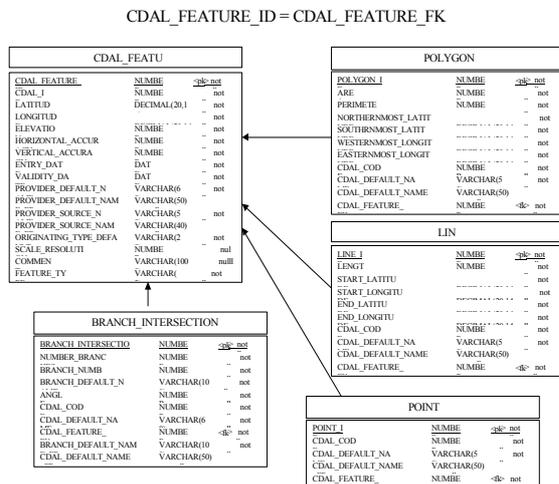


Figure 7: The CDAL data model for control points

Figure 7 illustrates the data model utilised to design the required tables and create the CDAL database. The proposed data model is in fact a modified version of the data model suggested by CTI (Centre for Topographic Information at Sherbrooke). The data are organised in a number of tables based on the type of the control point. In general, the CDAL data set consists of four classes of control points (Parent 2000):

- *Intersection Point (IN)*: Point is derived from an intersection feature in the source data set, its position corresponds to the point where at least three branches (roads, railways, highways, or access roads) meet. A branch is a linear feature that has one of its extremities at the intersection.
- *Point originating from a Point feature (PP)*: Point is derived from a point feature in the source data set. Its position corresponds exactly to that of the originating point.
- *Point originating from a Linear feature (PL)*: Point is derived from a linear feature in the source data set. Its position corresponds to the geometric centre of the originating line.
- *Point originating from a Polygon (PS)*: Point is derived from a polygonal feature in the source data set. Its position corresponds to the centre of mass of the originating polygon.

Accordingly, there are 4 tables (special tables) that store the information exclusive to a particular type of control point. These tables are Polygon, Line, Point, and Branch_Intersection. There is also one main table, CDAL_Feature that holds the common attributes of control points regardless of their type. This table is indexed on Latitude and Longitude fields.

The input text file contains data for a specific type of control point. Therefore, each run of the conversion program results in the insertion of records into the main table CDAL_Feature and one of the special tables (Point, Line, Branch_Intersection, and Polygon).

6. CONCLUSION

This paper discussed the design and implementation of a new online service that serves two purposes. The first permits searching of the CGDI Data Alignment Layer (CDAL), using a region of interest, through a Hypertext Transfer Protocol (HTTP) interface. That is, CDAL points are returned that fall in any bounding rectangle, specified by latitudes and longitudes. This service directly benefits the CGDI by providing a new CDAL access mechanism. The second service, built on the first, is a rendering service, which will visualize requested points of the CDAL. Such a service is of economic benefit for the production of geo-coded satellite/aerial imagery, where ground control is often expensive and difficult to obtain.

The operation concept developed in this project can be used in similar projects that aim at leveraging the power of the Internet to create on-line client/Server geo-spatial applications.

7. REFERENCE

Apache (2001), *Apache: HTTP Server Project*, <http://httpd.apache.org/>

CDAL (2001), *CDAL: A Service to Access and Visualize Data Alignment Layer*, <http://cdal.mda.ca/main.html>

CDAL (2000), *Creating the CGDI Data Alignment Layer (CDAL)*, <http://ccdi.icdg.gc.ca/html/main-e.html>

Parent, C. (2000), *Where are we with the CGDI Data Alignment Layer?*, GIS 2000, Toronto, Canada.

Parent, C. (1999), *The CGDA Data Alignment Layer*, Geomatica Vol. 53, No. 3.

Tomcat (2001), *The Jakarta Project: Tomcat*, <http://jakarta.apache.org/tomcat/index.html>

8. ACKNOWLEDGEMENTS

The authors express their gratitude to GeoConnections/NRCan (GeoInnovations program) for partial funding support of this project. RADARSAT International helped to define system requirements and assisted the project by continued evaluation of the system. We acknowledge the Center for Topographic Information-Sherbrooke for their support and providing the CDAL data set.