

Overcoming challenges of CAD conversion to GIS

Fathi Dwaik, Chandu Rao

Saudi Aramco
Dhahran, Saudi Arabia
fathi.dwaik@aramco.com

KEY WORDS: MGE, MGDM, ESRI, ArcSDE, Geodatabase

ABSTRACT:

Saudi Aramco's Surveying Service Division (SSD) provides a full range of hydrographic, land, geodetic and photogrammetric surveying and mapping services. In the last year, SSD has successfully completed the base mapping data migration from Intergraph MGDM database to the ESRI ArcSDE Geodatabase using custom cleanup and conversion tools. Over a hundred data layers, with numerous features covering of all Saudi Aramco's communities and reservations, are now GIS enabled to better serve the Company. This paper touched on Geodatabase model creation, data migration methodology, issues encountered along the way, and how they were resolved.

1. Introduction

Surveying Services Division (SSD) of Saudi Aramco is divided into three main operational areas: Geodetic & Photogrammetric Survey Section, Land Survey Section (LSS), and Hydrographic Survey Unit (HSU). The base mapping system employed at SSD consists, in general, of ImageStation, MicroStation, Modular GIS Environment (MGE) and Modular GeoData Manager (MGDM). Until June of this year, these modules were used to capture digital planimetric and topographic data, store captured data to MGDM database, and generate mapping and products for a large number of proponents within Saudi Aramco.

About a year and a half ago, SSD started the process of converting its MGE/MGDM-based map database system into a new data format based on ESRI GIS technology. Prior to this migration process, SSD commissioned Saudi Consolidated Engineering Company (SCEC) to build a conceptual Geodatabase design to carry out this conversion. The following section briefly describes the steps followed in building the SSD Geodatabase model.

2. Geodatabase Design

The project team conducted several interviews with internal and external user groups of SSD data to better understand their business processes and to identify their needs of geospatial information and GIS functionality. The outcome of these interviews, together with analysis of the existing MGDM data structure that supports the organization's functions was used to prepare a conceptual database design. MGDM mapping objects with common properties have been identified; the relationships between these objects were documented and modeled by Unified Modeling Language (UML) using Visio software. Proper geographic representation (points, line, polygon, annotation, and object without geometry) were assigned to the database objects.

The SSD MGDM base map is composed of Community and Reservation databases. The main difference is that Community features are captured and mapped at 1:1000 while Reservation features are captured and mapped at 1:10000. Both Community and Reservation databases were divided into major, minor, and detailed features. In developing the logical database model, those features were transformed into feature classes and subtypes and then grouped into different Feature Datasets in the new Geodatabase design (Transportation, Pipeline, Hydrography, Vegetation, Utilities, etc.).

Several issues impacting the spatial database and had to be considered at early stage of the Geodatabase design. Choosing appropriate map projection is important. Lambert Conformal Conic is used by SSD to store spatial data kingdom wide. Scale difference between Community and Reservation databases caused some duplication when real-world objects are represented by the same geometry type in both databases. Security measures to control access to the spatial database were considered at three different levels; at the operating system level where the user must have a username and password to login to the local system. The Relational Database Management System (RDBMS),

Oracle™ in this project, enforces users' roles to access spatial data set by database administrator. Access specific parts of the database could be controlled at feature class level by Geodatabase management tools of ArcCatalog application.

3. MGDM to Geodatabase Migration Process

3.1 Process Overview

Through examining the existing MGDM data for migration, a number of geometrical inaccuracies, as well as miss-coded features were identified. In many cases, the feature's symbology is not linked with its correct feature code. Therefore, all features were re-feature coded based on their symbology. The geometrical errors were mainly: open polygons, overlapping buildings, silver polygons, and duplicate vertices. Several custom tools were developed to automate the data cleanup process and to rectify these errors. A prototype project was designed to test the developed tools on two selected areas from both the Community and Reservation databases.

3.2 Methodology

The migration methodology incorporates mainly of the following processes:

- MGDM data conversion to Area DGN files
- Geometric cleanup of DGN files
- Convert clean files to Geodatabase
- Check the data quality of converted data

To better manage the data migration, the apparent seamless MGDM database was divided into DGN files based on area identifier. A custom script was created to extract the geographic extent of all CFT files. Each CFT file was assigned and AreaID based on CFT file location (COM/RES database) and geographic location.

The clean up process maintains the geometric integrity of the DGN files within defined tolerances, without breaking complex elements into simple elements. Additional tools and utilities were built to clean up various types of errors exist in MGDM database. For example, tools were developed to clean up elements that were not feature coded, or elements that cause geometric inconsistencies in the geodatabase. The following examples illustrate partially erroneous cases exist in MDGM and the measures taken to rectify them. Figure 1 shows an open polygon. This geometric inaccuracy is corrected automatically using cleanup custom tools.

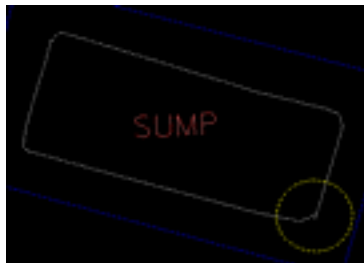


Figure 1: Open polygon

Figure 2 and 3 show self-intersecting polygon and duplicate vertices with different z-value.

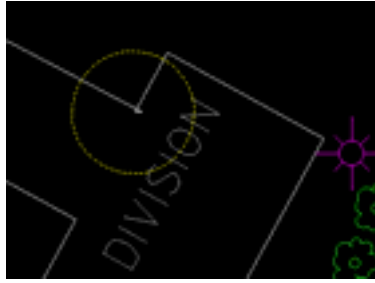


Figure 2: Self-intersecting polygon

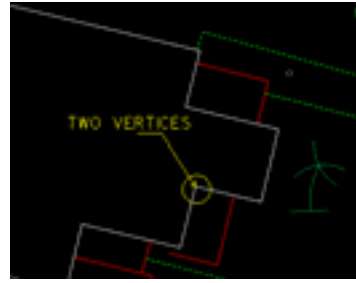


Figure 3: Duplicate vertices – different z values

Using fonts, rather than cells, as symbols, the origin of the MGDM font symbology is located at the center of the symbol. These symbols will be shifted to the new origin at the lower left corner of the element. Figure 4 illustrates the misplacement problem.



Figure 4: Misplaced symbols

Figure 5 shows storage area captured only by text annotation. The associated geometric feature is captured using surrounded features like fence, roads, building, etc., and/or reference photography as shown in Figure 6 below.

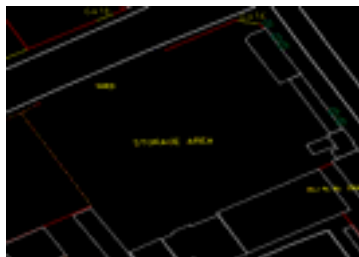


Figure 5: Area represented by annotation



Figure 6: Create polygon feature

The cleaned area DGN files are now ready to be converted to ESRI shape files. In addition to points, lines, and polygon conversion, text features are also converted to point features with width, height, and rotation attributes. The DBF files of the shape files were extended to create additional fields for all attributes available in the MGDM database. A custom script was developed to extract features attributes from MGDM Oracle™ database and associate them with the corresponding DBF shape features. To import the shape files into Geodatabase, a bulk import utility was developed to run inside ArcGIS. This utility creates a log file to report any problem during the import process, and the total number of features loaded.

The last process is to ensure completion of conversion and quality of the data converted. To achieve that, a number of quality control tools and procedures were developed to check each part of the conversion process and ensure that no elements or attributes are lost. The number of elements of each type in the original CFT files was compared with the number of elements after the merge process to create the area wide DGN files. Prior to any modification to the area DGN files, a program is used to extract all attribute linkages. The output log is compared with a second run of this

program prior to the DGN conversion to shape file. The same program, along with a separate program to extract attributes based on MSLink, are used to identify orphan records and graphics. Orphan record is one exist in the database where its associated Microstation element no longer exists. Orphan graphics, on the other hand, is an element with a link to a database record; this record no longer exists. Figure 7 summarizes the process followed and the tools used for migrating the database from MGDM to Geodatabase.

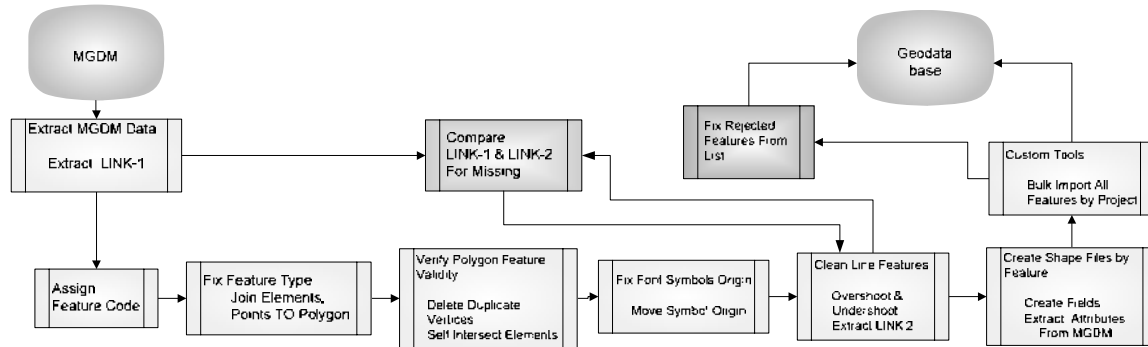


Figure 7: MGDM to GDB migration workflow

3.3 More Tools to Overcome Challenges

In addition to the cleanup and migration tools described above, more custom tools were developed to resolve specific problems. For example, remove small loops usually created erroneously when digitizing vertical lines. Zero length lines could be identified and deleted automatically or visited one at a time for verification. In addition, programs were used to resymbolize elements based on their feature code; level, weight, color, and line style are corrected based on element's attached feature code. Note that the symbols are not imported to the geodatabase, but correct symbols help the project team to interpret DGN files more accurately.

More cleanup tools were used to check for multiple links to database records; in such cases, elements are visited to divide the element or create a duplicate one. The remove graphics cells tool creates selection set of all complex elements (cells that are not shapes or simple cells). The tool drives to each element in the selection for verification and drop of the element. Otherwise all complex elements can be dropped automatically.

4. Conclusion

Migration spatial database to Geodatabase format offers several advantages to SSD internal operations and well as external customers. In addition to seamless physical storage of GIS data inside RDBMS, Geodatabase data model supports multi edits to the same database. The database conversion helped to standardize the spatial data formats across several departments within the corporate.

The MGDM database was migrated to new data model that is transitional, in the sense of lacking complete spatial relationships and comprehensive topology rules. As SSD continues to develop and populate the geodatabase, a growing interest to use the spatial data to deploy GIS applications that adds efficiency and effectiveness to the corporate workforce.

5. Acknowledgement

The authors wish to thank Amer Dulaijan and Mohammad Khatib for their significant contribution and comments reviewing this publication.

6. References

MAPS Geosystems: *MGDM to Geodatabase Conversion Project # 400303/041 – System Documentation*. Draft v2.3, 2004, Surveying Services Division, Saudi Aramco.