Australian Journal of Basic and Applied Sciences, 8(2) February 2014, Pages: 99-105

AENSI Journals

Australian Journal of Basic and Applied Sciences

ISSN:1991-8178

Journal home page: www.ajbasweb.com

Layers Resolving Problems of Conversion from CAD files to Geographic Information System

Dr. Samih Al Rawashdeh

Associate Professor at Al-Balqa' Applied University (BAU) Jordan, SALT (19117), Faculty of Engineering, Department of Survey and Geomatics Engineering

ARTICLE INFO

Article history:
Received 23 December 2013
Received in revised form 25 February 2014
Accepted 26 February 2014
Available online 15 March 2014

Keywords:
CAD, GIS, Topology, Shape Files, Geodatabase.

ABSTRACT

There are great problems of conversion from CAD files to geographic information system (GIS) due mainly to that CAD drawings and GIS data are created for different purposes, and as a result the software's data models are inherently different, and the files differ in their construction and attributes. CAD data contain information to draw it, with features organized by mixed layers. While, GIS data has an additional functionality; its feature can hold enormous amounts of data describing the feature and organized in one feature type layers. Finally, GIS data is "spatially informed" regarding adjacency and other spatial relationships. Then the conversion of CAD files to shape file will face problems and the data will be corrupted. The main objective of this paper is to establish a methodology using a case study to provide solutions using different software and programs to resolve the problem of conversion from Cad to GIS. This allowed the reduction of the problem of this type of conversion.

To Cite This Article: Dr. Samih Al Rawashdeh., Layers Resolving Problems of Conversion from CAD files to Geographic Information System. ***Aust. J. Basic & Appl. Sci.,*** *8(2): 99-105, 2014*

INTRODUCTION

Increasing demands for the data and information of a spatial nature in the last three decades, led to the vast development of automated tools for efficient storage, analysis and presentation of geographic data. This rapidly evolving technology has come to be known as "Geographic Information Systems (GIS)". Geographic information system goes beyond description; it also includes analysis, modeling, and prediction. According to the Environmental

Systems Research Institute (ESRI), a GIS is defined as "an organized collection of computer hardware, application software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographic referenced information.

Geographically referenced data separates GIS from other information systems. For example to describe a road in GIS, we refer to its location (i.e. where it is located?) and to its characteristics (length, name, speed limit etc.). The location, also called geometry or shape, represents spatial data, whereas characteristics are attribute data. Thus a geographically referenced data has two components: spatial data and attribute data.

Recent advances in GIS software have made it easier to transfer CAD drawings into GIS shape files or geodatabase, but many quirks still exist in file transfer due to the inherent differences in the file formats and structure. With this project, the NPS developed a step-by-step method of transfer, noted issues in both the CAD and GIS to be aware of when attempting to convert CAD files into GIS format, and created GIS-friendly guidelines for outsourcing new CAD work. This document is intended to be a guide, and as such cannot address every technical issue and possible scenario in CAD to GIS transfers. But still no sufficient solution to be compatible to GIS.

The applications of Geographic Information system (GIS) grow rapidly to satisfy the need of GIS operators. Since GIS import data from many types of sources, this presents a great challenge because these sources have structures of data differ from GIS data structure. The most important source of data for GIS is the Computer-aided design or Computer Aided Drafting (CAD) files. So, importing data from these sources require some necessary treatment and management such as building topology, modifying the attributes, editing the spatial features, separating and classifying of layers. Recent advances in GIS software have made it easier to transfer data into GIS shapefiles or geodatabases, but many quirks still exist in file transfer due to the inherent differences in the file formats and we suffer the scarcity of works which deals with the data transfer compatibility, integrity and accuracy.

Corresponding Author: Dr. Samih Al Rawashdeh, Associate Professor at Al-Balqa' Applied University (BAU) Jordan, SALT (19117), Faculty of Engineering, Department of Survey and Geomatics Engineering.
E-mail: rsamih@hotmail.com.

Data migration is a labor-intensive process and often is the most expensive step in the entire GIS implementation. The key to data migration is the data accuracy. In order to produce accurate data, the data conversion vendor should be provided a data conversion specification generated for the data model before starting full production. Data validations in both CAD and GIS are critical to the success of data migration. Data validation routines need to be developed before starting full production. These routines are typically developed in CAD using Visual Basic *(*VBA) and in GIS using Acute Myeloid Leukemia (AML).

The growing interest in applying Indoor GIS technology is mainly to facilitate management applications which have now led BIM, CAD and GIS practitioners into a realm. This necessitates a clear and defined passage from legacy and new design BIM & CAD data to a fully constructed GIS Geodatabase server environment. Xiaodong (2001) built a Geodatabase into a Multi-user GIS, migrating CAD Data to a Geodatabase, and GIS. Linyan *et al* (2011) clarified how to overcome some types of problems in the conversion of CAD data in to GIS for land planning purposes. They outlined a methodology which produces the desired end result: GIS integration with Building Automation Systems, Computerized Maintenance Management Systems, Energy Management Systems and Enterprise Asset Management. This commingling of existing BIM and CAD data with robust GIS systems and techniques has created a new paradigm that offers facility managers, law enforcement and security personnel, plant operations managers and space utilization managers unprecedented access to disparate and normalized data resources: A BIM and CAD to GIS workflow that creates a new data source that the geospatial world can utilize for the purposes of Indoor Positioning, RFID, RTLS and Indoor Routing Applications.

The problem of editing was resolved in most GIS software, the Geodatabase, through ArcSDE in ArcGIS, supports multi-user access and editing, versioning and rollback. Relying on a relational database structure, Geodatabase allows users to define objects and behaviors (Zeiler, 1999).

A report conducted by the company ESRI –Infograph (2005), emphasized the importance of topology for the editing process. This document is intended to be a guide but it didn't address every technical issue and possible scenario in CAD to GIS transfers and Al Rawashdeh in 2008, processed and interpreted the problem of data importing or converting and their effect on the quality and the integrity of the data.

The main objective of this work is to put a methodology for resolving the unsolved problems of conversion from CAD files to GIS. For that a real application of conversion from CAD files to GIS files will be done and the site of Al-Balqa Applied University is chosen to be the zone of application.

Problems of data Conversion:

CAD and GIS data are created for different purposes, and as a result the software's data models are inherently different, and the files differ in their construction and attributes.

CAD file deals with points, lines, or polylines and does not contain additional attributes. It contains only the information needed to draw itself, such as line weight (thickness), line type (continuous, dotted, and dashed), color, and the layer to which it is assigned. GIS data, on the other hand, has an additional functionality; its feature can hold enormous amounts of data describing the features (a polygon representing a house could contain information about the owner, street address, numbers of bathrooms, bedrooms, etc). In addition, GIS data is "spatially informed" regarding adjacency and other spatial relationships. Other main difference in CAD data is that features are organized by mixed layers, which can contain points, lines, polygons, and annotation as compared with a GIS shape file, which can only contain one feature type (each drawing feature in GIS belongs to a specific layer). In this way, elements of similar type can be viewed, hidden, frozen, moved, and edited together. An example is a layer named “TREE” that contains all the elements representing trees in the CAD file. When a CAD drawing is exported to a GIS shape file, the CAD drawing's layers become a field in the GIS shape files attribute table. New point, line, and polygon shape files can then be extracted based on each layer (trees, buildings, roads, etc). These issues will be addressed in our step-by-step methodology for the conversion of CAD to GIS

Methodology:

Software and programs:

Many software and programs were used in the processing such as: Arc GIS, Geomedia Professional, AutoCad, ASP.Net and ETGeoWizards, each for a specific reason.

Data Collection and application:

A GIS and a database is built for the BAU using field collected data about the personnel in the buildings, the location of their offices within the building and the existing facilities. The plans for the buildings floor by floor in CAD files were used to locate the offices within the buildings. Figure 1, presents a part of the study area in raster and in CAD files and Figure 2 presents samples of the plan in CAD Drawing files for some floors of the engineering college.

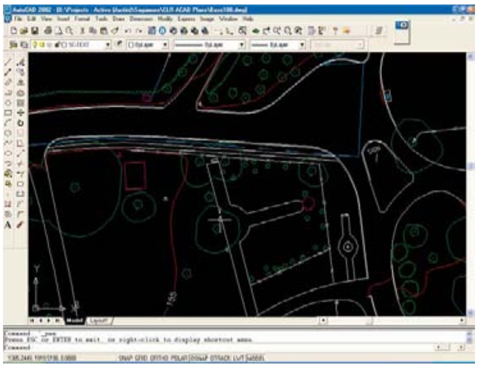

Fig. 1: A part of Al-Balqa Applied University drawn in CAD files.

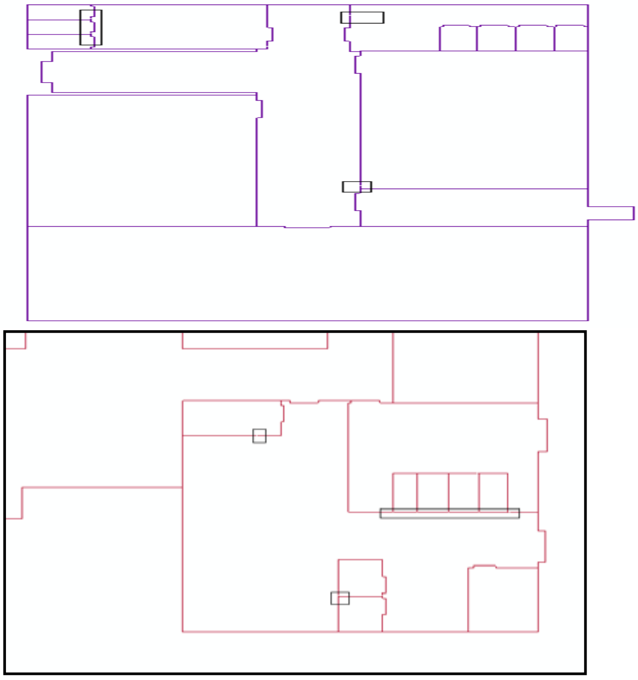

Fig. 2: Floor plans (CAD files) for the engineering building.

2) Creating GIS Shape files from CAD layers:

When a CAD drawing is imported into GIS, it will interpret all elements into the four primitive geometric GIS features: points, lines, polygons, in addition to annotation then, in to shape files or feature classes. This division is based on the original CAD layers. There are no interaction between data and no connection to the real world or to other design objects.

One of the most popular CAD packages across all industry sectors is Autodesk's AutoCAD, but joining AutoCAD designs and corporate mapping databases is currently not a simple task. To achieve this, an AutoCAD user has to find some way of translating mapping data into DXF (AutoCAD format) files, and then storing these files somewhere for use. During the translation process, the source data may have been updated in a way that can affect the accurate depiction of the design.

Conventionally, GIS contain two types of data: (i) raster data (two-dimensional arrays of data values), such as those generated by satellites and other remote sensing operations; and (ii) 'vector' data (lines and polygons).
The method of conversion we used begins from the transfer of CAD files to GIS using:

1. Geomedia software,
2. ArcGIS software and,
3. CAD 2 shape program

The use of Geomedia Professional software:

This product allows combining geographic data from different sources, in different formats, and with different map projections, all into a single environment. It allows also performing complex queries on spatial

and attributes data from various sources, and produces numerous views of highly sophisticated maps in a single GeoWorkspace. The applicability of GIS project directly depends on the accuracy and topological consistency of the vector data that form the basis for GIS. The existence of errors or inconsistency in the vector data could easily prevent a reliable GIS analysis.(Bo, 2008 ; Linyan, 2011) Therefore, it is necessary to do an overall check as well as to compare them in terms of integrity in order to eliminate the errors coming from the encoding process. Using different software allows better debugging and error detecting than using one GIS software.
The application is done as follows:

All the AutoCAD drawings for all the features in the study area (for example: building, streets) were transferred to GIS feature classes or shapefiles. This required the following operations:

1. All closed polylines should be converted to polygons.
2. creating attributes for the transfered polylines
3. For advanced polygon creation we have to use the Build Polygon Wizard.
4. Separating the layers depending upon their geometric feature.
5. Editing process which will involve intercepting errors and modifying features; the attribute data will be built and will be transferred to the corresponding polygon features (linked to their spatial features).

3) Using Topological data structure to resolve problems of conversion:

Topology examines the characteristics such as neighborhood and merges outside the coordinate system information of geometric objects (Chung *et al* 1995; Chung *et al* 1997)

Topology is the geometric relationship between edges, nodes and the faces. According to the other definition, topology is a way or method in which logical relations can be defined such as neighborhood, coincidence, inclusion, intersection, sharing, in addition to metric relationships such as the geometrically identifiable coordinate, length, area (Chang, 2002).

The aim of the topological knowledge in GIS is to increase spatial analysis opportunities. To represent the spatial information as well as spatial relations (neighborhood, coincidence, directions, links) of the features in topological data structures on the computer; the node elements corresponding to point, edge (arc) elements corresponding to the lines and the face elements corresponding to the polygons are used.

Note that more than two line intersection points may be an exposed end of the line or a single point not connected to any side. Edge is a set of coordinate pairs starting with a node and ending with a node. Face is the largest two-dimensional space restricted by the edges and cannot be divided by an edge. According to this, point features are composed of nodes, line features with one or more edges, and the face features surrounded by one or more edges.

To be able to evaluate a topological database, in addition to the geometric properties, the following relationships must be determined and stored:

a) Edges making up the boundaries of each polygon
b) Neighborhood relations between the polygons
c) Connections at the intersection points
d) Start and end points of edges

Accurate topological data leads to correct and reliable spatial analysis, in addition to cost reduction. In the databases where topological properties can be stored and used, information about a line feature is connected to another line feature, depending on where, and at what point line features are combined, the polygon features to the right and left line features can be easily and quickly queried.

The existence of the topological data in addition to the spatial data and attribute data in database will cause the increase in volume of the database. The advantage of topological data that allows spatial analysis and queries expected from GIS offers very significant advantage compared to data storage disadvantages.

Topological errors in spatial vector data and its visualization and elimination by means of AutoCAD Ve auto lisp. Since spatial vector data contain many topological errors arising from the quality of the source material, data collection techniques and the human interaction, acquisition of vector data for GIS requires further processing in order to use it for some advanced spatial analysis. Before using spatial data for analysis, it must be made topologically-correct. Vector features can be made to respect spatial integrity through the application of topology rules such as; ‘’polygons must not overlap’’.

To make data usable or to make it clean, some quality checks should be applied. All GIS software has the functionality of performing this kind of processes. Most of the GIS software is also capable of providing the spatial integrity and cleanness during data capturing. The topological errors arising from the conversion into the digital maps could be easily removed through the software written in AutoCAD LISP.

Australian Journal of Basic and Applied Sciences, 8(2) February 2014, Pages: 99-105

RESULTS AND DISCUSSION

Adding CAD files to a Dataview:

CAD drawings can be immediately displayed in ArcGIS. This step does not create GIS data; it only displays CAD data in the GIS dataview. It should be converted to GIS layers.

Exporting CAD data as Shapefiles:

When CAD file is added to GIS software, it will be divided into data sets, based on the four types of GIS data: point, line, polygon, and annotation. It is to be noted that exported CAD files will be in two types of files. The annotation set cannot be exported as a shapefile. Annotation must be exported as a feature class of a geodatabase. The exported files from CAD to GIS layers will be separated in to layers having the same primitive geometric features: points, polylines or polygons.

To repair specific topological errors in the vector data of our application such as:

1) Floating or short lines
2) Overlapping lines
3) Overshoots and undershoots
4) Unclosed and weird polygons

Integrity of Data coming from CAD files to GIS for the application purposes:

In order for newly created shapefiles to be compatible properly with other GIS data, they need to have a coordinate system projection defined in GIS. The Universal Transverse Mercator (UTM) coordinate system was chosen the standard datum will be NAD83 for all layers in this application. To determine if a shape file has a defined coordinate system, we have to check the original CAD drawing notes, compare coordinates of an identical spot in CAD and GIS, and/or check with the CAD drawing developer. Another method to determine the coordinate system consists of adding the file to a blank document without any information and then we give a system of projection to this layer. Some layers are given a system of projection but the overlay with other layers doesn't properly occur (overlaid). In this case, it requires 4 correctly distributed points (at least) to be defined in x and y coordinates, then we have to perform the geometric correction. This process will move the shapefile from its current coordinates to the correct coordinates.

Another problem we faced was that the original CAD drawings were created using a defined coordinate system not recognized by the GIS software. In this case we specified the parameters in the "Define Projection Wizard" and then we had to spatially adjust the shapefile by using the similarity or the affine transformation, this depended on the type of errors. The Similarity Transformation maintains the aspect ratio because it does not differentially scale or skew the obtained data. If this algorithm does not perform well, we have to choose "Affine transformation". However, the affine transformation will differentially scale, skew and rotate translated data. The remaining adjustment methods were often too aggressive to be useful for CAD data adjustments. In this way, the spatial adjustment can be executed and we have to complete the adjustment. When a satisfactory match has been achieved, we have to stop the Editing process. The shapefiles are now properly overlaid with other UTM NAD83 data. The obtained residuals were satisfied (0, 0012 meter).

Cleaning CAD layers from floating or short lines:

Before transferring the CAD drawing to GIS, we have to verify that all features are on their respective layers. We have to isolate each layer by turning all other layers OFF and to see that there are no stray drawing elements or features that belong on other layers. The elements which does not belong to the layer must be removed (or to cut and to paste) into their proper layer. Layer 0 and "defpoints" are inherent in all CAD drawings which cannot be deleted, so they should contain no objects.

Removing duplicating arcs:

According to topological integrity rules, lines must not overlap and duplicating lines must be removed to avoid obtaining false results when we do spatial analysis. For this operation, the relevant layer should be chosen, the duplicate arc could be optionally moved to a new layer (Fig. 3) to be deleted.

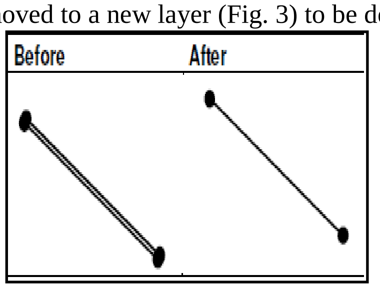

Fig. 3: An example of removing a duplicated layer.

Correcting the errors of overshoot and under shoot:

The errors of overshoot and undershoot create a great problem for GIS analysis and give fault results for networking process. In finding the shortest distances between features, they will not give the real shortest distance especially in the case of the undershoot errors. To correct this type of errors we use the editing process in GIS (Fig. 4).

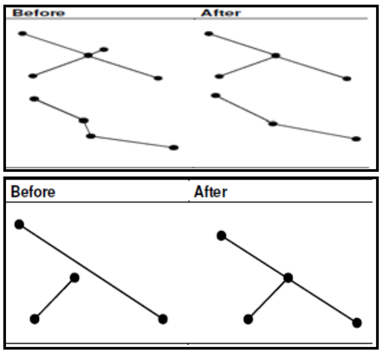

Fig. 4: Correcting undershoot and overshoot errors.

Correcting opened polygons:

This type of errors forms a great challenge for the correct conversion of layers from CAD to GIS layers, considering these polygons polylines instead polygons. Many errors in the data base were marked (Fig.5).

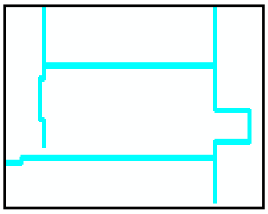

Fig. 5: Unclosed polygon detected.

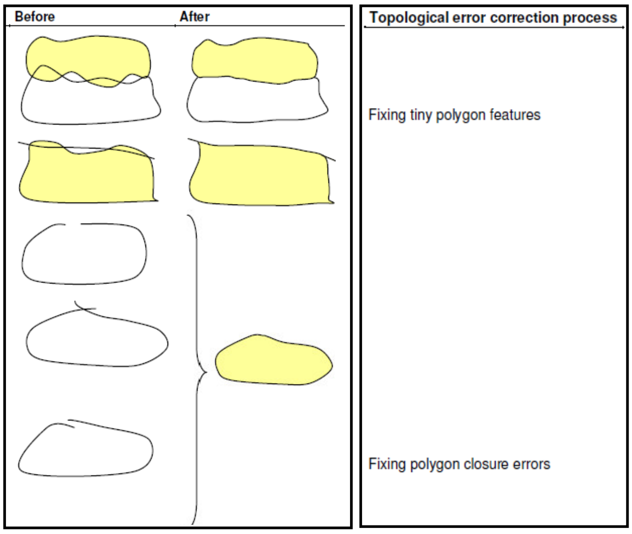

Fig. 6: Topological errors.

Create new shapefiles from CAD files using the attribute table:

When a CAD drawing is converted to a GIS shapefile, the original CAD layers become a field in the shapefile’s attribute table. By selecting objects in the shapefile based on their CAD layer names, and exporting them as new shapefiles, shapefiles can be created for each original CAD layer. Therefore, this procedure can yield separate shapefiles for grouping objects (trees, buildings, roads, etc) which can then be integrated more effectively into a GIS database.

Correcting topological error in our application:

This step is very important for obtaining a reliable and correct data base. It modifies and corrects all the topological error which can happen in the geodatabase. The most important types of topological errors consist of neighborhood, adjacency and relationships between different data base. Two polygons shared an arc must be corrected from the sliver or splinters errors (Fig. 6). We note that splinters or slivers are thin polygons between two adjacent polygons due to encoding errors and must be eliminated to obtain a reliable, complete and correct database.

Adding attribute data:

CAD files are not very efficient to link non spatial data to their spatial data then, for all features we have to build attribute data. In addition, GIS data must be organized and managed in different ways in the data base. For example: there are no relate tables in CAD files and we haven't the capability to do data fusion or other operations of data management as in GIS software.

Conclusion:

The process of integrating different types of data into GIS is very important these days as different types of software do different functions more practically and accurately. More and more functions and capabilities and more friendly to users software necessitates the conversion of the data while keeping its integrity. This work outlines a methodology for the transfer from CAD files (butter drawing and 3D modeling capabilities) to GIS database. This requires processing and editing to build complete, efficient and correct database. This is important to allow the full functionality to perform the different operations of overlay, analysis network or any other GIS operations. These necessary operations are due to the nature of the two types of software. CAD Autodesk offers integrated CAD and GIS solutions for surveying, mapping, civil engineering, and infrastructure lifecycle management. These solutions create, integrate, and manage CAD and GIS features with no loss of precision or topology.

REFERENCES

Al Rawashdeh, S., 2008. Effect of data processing in data quality, Journal of computer science, Science Publication, 4(12): 8.

Bo, S., Y. He, X. He, N. Zhang, Z. Li, 2008. "Study on Interactive Editing Method in 3D GIS Environment," *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*, pp: 1-5, 12-14 Oct. 2008doi: 10.1109/ WiCom.2008.2979.

Linyan He, Guoping Wu, Dongmei Dai, Li Chen, Gongxun Chen, 2011. "Data conversion between CAD and GIS in land planning," *Geoinformatics, 2011 19th International Conference on*, pp: 1-4, 24-26. June 2011 doi: 10.1109/GeoInformatics.2011.5981193 -URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=& arnumber=5981193 &isnumber=5980662.

Chang, K., 2002. "Introduction to Geographic Information Systems", Tata McGraw-Hill, New Delhi.

Chung, M., I. Jhoty, F.M. Li Yim and J. Deville, 1995. GIS in Mauritius. 10 p. In: Conference Proceedings of Environmental System Research Institute (South. Asia) 4th Annual User Group, 4 - 6 September 1995. Singapore: ESRI, 8 pages.

Chung, M., G. Pillay, R. Domaingue and I. Jhoty, 1997. "GEOGRAPHIC INFORMATION SYSTEM - A TOOL FOR SMALL CANE GROWERS OF MAURITIUS, AMAS"., Food and Agricultural Research Council, Réduit, Mauritius online at: http://www.gov.mu/portal/sites/ncb/moa/farc/amas97/pdf/msiri09.pdf

ESRI -Infograph Company, 2005, GIS Topography, an ESRI white paper, 14 pages, online at (accessed on January 2012):http://www.esri.com/library/whitepapers/pdfs/gis_topology.pdf

Xiaodong, H., 2001. "A ROAD MAP TO IMPLEMENTING AN ENTERPRISE GIS", Proceeding ESRI, online accessed on January 2012, http://proceedings.esri.com/library/userconf/proc01/professional/ papers/pap315/p315.htm.

Zeiler, M., 1999. Modeling Our World – The Esri Guide to Geodatabase Design, ESRI Press.