PSIG 1105

Maps To Models - Building Distribution System Models From GIS or CAD Data

Brad Bean, Partner - Bradley B Bean, PE

This paper was prepared for presentation at the PSIG Annual Meeting held in Napa Valley, California , 24 May – 27 May 2011.

ABSTRACT

Geographic Information Systems (GIS) and Computer Aided Drawing (CAD) applications are used by many Natural Gas Utilities to maintain information about the facilities in their gas distribution systems. This data is often the same or similar to the information required to build and maintain a hydraulic model of their gas system. This paper will discuss and identify several key issues that should be considered to allow efficient exchange of data between GIS/CAD and modeling applications.

A discussion of the difference between GIS and Computer Aided Drawing (CAD) applications will be presented. Common generic exchange file formats will be described for each application type. Data related topics such as coordinate systems, facility identification and separation, pipe size designation, connectivity, customer load allocation, regulator and gate station handling, inclusion of valves, and the use of operating data will be discussed along with common issues associated with each topic. A list of common problems encountered, and suggested considerations will be presented.

Although focused primarily on building models for distribution system applications, the topics in this discussion apply to all piping systems types.

CAD-GIS HISTORY

Modern electronic mapping and drawing systems began to spawn during the 1980's. Computer-Aided-Drafting (CAD) systems began to be used by natural gas operators to map their facilities in the early part of the decade. Systems such as AutoCad (1982) and MicroStation (1985) were PC based applications that provided simple two dimensional functions suitable for drawing and mapping line work. This technology allowed utility operators to move from purely paper based record systems to electronic based systems.

Later in the decade, Automated Mapping / Facilities Management (AM/FM) systems began to appear. At the time, these systems were often referred to as “smart maps” and were the predecessors to what we know as GIS now. These systems merged CAD and database technology allowing mapping and facility data to be maintained by the same application. During this time systems were available from vendors such as Intergraph, McDonald-Douglas, and IBM. These platforms were expensive, often required specialized hardware to operate, and thus had a limited user base.

Towards the end of the decade, Geographic Information Systems (GIS) began to be applied to utility network type systems. Prior to this time, GIS was mainly used in traditional natural resource fields, such as forest and wildlife management. GIS was largely used to manage area type features, such as tree species or soil type in a defined area. These systems relied heavily on area geometry like polygons to manage and identify the physical features that they worked with. Managing linear features, like roads and utility networks was not well supported by these systems in the beginning.

AM/FM and GIS systems were similar, but with not so well defined differences. AM/FM systems were generally based on a graphics or mapping system that connected to a separate database to access facility data. The graphics data and attribute data were kept separately. GIS systems were generally based on a database system that used a graphics system to display or map the various geographic features contained in the database. More or less the same thing in terms of integration, they both used two separate systems to maintain the two different types of data - graphic and attribute. In a general sense, AM/FM systems tended to be more focused on managing facility data, where GIS systems tended to be more focused on analyzing data and visualizing the results.

As technology advanced, 2D CAD systems continued to evolve into highly sophisticated 3D graphics systems, and the AM/FM and GIS systems essentially evolved and merged into the technology known singularly as GIS.

NETWORK MODELING HISTORY

By the early 1980's computerized network modeling had been in existence for some time. It was mainly implemented on “main” frame computer systems using text based software with little or no graphic representation or display of the data or results. In the early days the user might enter data by creating a set of punch cards or typing data using a “terminal”, then send the data for processing. In later systems the user might create an ASCII text file that would be processed by the modeling software. Results would usually be provided in the tabular printed report.

Although these systems were attempting to model the same facilities that were being captured by the newly emerging CAD, AM/FM, and GIS graphic and database systems, there was little or no interaction between the technologies. Anyone doing modeling back in the day, probably remembers having a paper map with the piping system drawn on it. The map would have pipe size, lengths, node names, and maybe pressures and loads or flows shown on it. It would invariably have notes and revisions scribbled all over it. This “node map” was the graphic depiction of the model. Although the everyday user might have a mental map of the model, without the node map it was very cumbersome for a new user to figure out how a model was put together and even more difficult to interperate the results.

With the introduction of the IBM PC and Microsoft Windows, a new generation of modeling software emerged. In the late 1980's Gregg Engineering introduced a Windows product which allowed graphical interaction between the model data and the user. Others were introduced in the years that followed. These new graphical based modeling systems made it more desirable to share the data that was in the CAD and AM/FM-GIS systems within the modeling application.

At the time it seemed only logical that these systems should “talk” to each other, and maybe even evolve into a single technology. Some attempts were made to integrate the two technologies. In the early 1990's, Bradley Bean offered a network modeling application that ran directly from within AutoCad. The product was well received by users, however there were issues with data management, users would often forgot to use a copy of their data to perform what if network changes, and would corrupt their actual mapping data. For this and various other reasons the product was eventually discontinued.

For as much as it seemed like a logical marriage, true integration of CAD-GIS systems and network modeling has not occurred. Today’s modeling and CAD-GIS systems do very well at what they are intended to do, but are truly separate systems - this is probably the way it will remain, at least for the near term. So with that reality, it is necessary and desirable to implement methods that allow data to be shared or more specifically be transferred between the technologies.

INTERCHANGE FILES

Today data is preliminarily shared between the modeling and CAD-GIS technologies by translating interchange files. In the pipeline industry several “standard” interchange formats have been attempted including the Pipeline Open Data Standard (PODS) format supported by the PODS Association, and the Extensible Pipeline Simulation Language (XPSL) format proposed by the Pipeline Simulation Interest Group (PSIG). Both serve a purpose but neither are really conducive to sharing data between distribution system modeling packages and mainstream CAD-GIS systems. In the CAD industry, the Drawing Exchange Format (DXF) has more or less been adopted as the generic interchange file for sharing data between purely mapping or drawing systems and other applications. In the GIS industry, the shape file format (SHP) has largely been adopted as the generic interchange format for sharing data with other applications. Other formats exist and are emerging, however for the purposes of this document, the DXF and SHP file formats will be the only formats explored in more detail.

Neither the DXF or SHP file types are particularly efficient in terms of file size, however this lack of compression eliminates the need to deal with sophisticated storage algorithms, and makes them very easy and efficient to read and write data to and from. With today’s computer storage capacity and performance levels, the inefficient storage characteristic is not really much of a consideration.

DXF Files

An “entity” is the basic feature used to define the map or drawing in a DXF CAD file. For example an entity might represent a single line, an arc, or a polyline.

A DXF file is a “coded” sequence of values stored in ASCII text format. It contains a list of code and value pairs. The code indicates the type of value that follows it. An entity is described in the DXF by a sequence of these code/value pairs. For example a sequence might consist of a code indicating the type of entity that follows, then be followed by a series of code/value pairs describing the coordinates associated with the entity.

One common feature of a CAD system is that it allows different entities to be placed or grouped on separate “layers”. Originally a layer was thought of as separate sheets of transparent paper that could be overlaid together to produce different views of the map or drawing. Nowadays layers might be thought of as individual folders that contain specific data features or entities in this case.

When used effectively, layers make working with CAD data very efficient. For example, all 2" Polyethylene (PE) pipe might be contained on an individual layer, or all pipe associated with a certain pressure tier, or all pipe installed during a certain year might be contained on separate layers. The best case is when layers are used to highly segregate the data, for example a single layer might contain all of the 2" PE pipe, associated with the intermediate pressure system, installed in 1999. Line type and color can also be used to segregate facilities in a CAD file.

A DXF file generally contains only graphical information, however a special implementation of the file can also contain non-graphical attribute information using special codes and specifically designed applications. For example a special code might be used to indicate a pipe size. Except for custom implementations, most network modeling systems cannot effectively work with non-graphical attribute data contained in a DXF file.

The DXF file format was originally developed and published by AutoDesk. Although it is vendor specific, it has been adopted in various forms by many CAD type systems.

Shape Files

In reality there is no such thing as _a_ "shape file". A shape file is actually a collection of files that represent various types of data. The most basic collection consists of a "shape" file which contains the geographic data - the coordinates - describing the "shape" of each feature. A "database" file which contains the various attribute information for each feature - for example pipe size and material. And an "index" file which links the shape file and database file together. This file configuration is not unlike the methods used by early AM/FM systems.

GIS systems also use a data grouping method similar to layers in a CAD file, in the shape file context, these groups are referred to as "themes". A theme can contain only one geographic type. A shape file can contain only one theme and consequently only one geographic feature type. For example a shape file can contain only linear features such as lines (arcs) representing pipes, or points such as customer locations, it cannot contain both. Through its indexed files, it can contain both graphic and attribute information. For example it can contain both location coordinate values, and pipe sizes.

The shape and index files are in binary format, the database file is in dBase III (3) format. The shape file format was developed and published by Environmental Research Systems Institute (ESRI). Similar to the DXF format, it is supported by many applications.

COORDINATE SYSTEMS

At the graphics heart of any CAD-GIS is the implementation of some sort of coordinate system. In the basic sense, the coordinate system allows the feature locations to be described by a series of numeric pair or triple of values.

CAD systems have traditionally used an X-Y-Z planar coordinate system. The X and Y values basically represent directions as they would be drawn on a flat sheet of paper, X representing left to right (west to east) and Y representing bottom to top (south to north). The Z direction represents elevation.

GIS systems have traditionally used a global (latitude and longitude) coordinate system. Latitude represents rings or parallels around the equatorial circumference of the earth. Longitude represents lines connecting the poles of the earth. Each are measured in degrees, minutes, and seconds. Elevation is just elevation.

Most network modeling systems require that feature coordinates be represented by planar X-Y values. This type of coordinate system is appropriate for representing comparatively "small" areas, however issues arise when trying to represent "large" areas. If you try to wrap a sheet of paper around a desk top globe of the earth, you will quickly see that it is not possible to get the paper to lay flat without creating "wrinkles" or without cutting the sheet into small slices. This is the main issue encountered when trying to use a planar coordinate system to represent large portions of the earth.

For large global areas, Latitude-Longitude (Lat-Lon) systems provide a good method of describing locations. However, Lat-Lon systems create a challenge in representing a continuous "map" on a flat sheet of paper or computer monitor. Some type of "projection" is required to flatten out coordinates expressed in these coordinate values. There are a number of projection styles and systems available.

In the US, the State Plane coordinate system provides one commonly used projection. The State Plane system is made up of zones more or less specific to individual states. Each zone has a certain origin and scale associated with it, and applies only to a certain region of latitude and longitude. Unfortunately, in some states, multiple zones exists. The challenge for the user, especially if they are modeling a system that covers a large area which may cross several State Plane zones, is what projection to use. Fortunately, network modeling does not require the accuracy of actual land surveying applications, so a compromise projection is often acceptable. For example choose the zone in which the propensity of the system being modeled resides, or just choose one or the other and be consistent.

Although network modeling does not require land surveying level of accuracy, the selection of a coordinate system is not trivial. The most important thing is that a *consistent* coordinate system be used. Using a consistent coordinate system will allow future data to be easily added to past models, and will allow overlay of other graphic data such as base maps depicting roadways, lot lines, and other land or topographic features.

EXPORTING DATA FROM THE CAD-GIS SYSTEM

In general most CAD-GIS systems do not normally work directly with their data in DXF or SHP file format. For efficiency and performance issues they usually use some type of proprietary file format. When needed, the DXF or SHP interchange file usually needs to be created by exporting data from the associated application.

CAD systems often use an internal or user define coordinate system with a shifted origin. When exporting to the interchange file, the output coordinate system needs to be considered and accommodated. For example Microstation uses a unique internal coordinate system, if not correctly specified during export, the coordinate values in the resulting DXF file will be of a very odd magnitude. In AutoCad, if the drawing units are set to Architectural or Engineering units, the coordinate values in the resulting DXF file will be "inches". This situation needs to be considered and addressed, sometimes requiring some processing of the data before importing it into the modeling application.

If locations in a GIS are maintained in Lat-Lon values, or in a projection system different than the one used in the modeling application, it will be necessary to specify the desired projection during the export process, or the exported data will need to be processed with a third party application before importing it to create the model.

One of the benefits of creating a model from "mapped" data is that the resulting model should be "to scale" eliminating the need to be concerned about pipe lengths. Care must be taken to ensure that the exported output coordinates are appropriate, this will ensure that the pipe lengths calculated using these values are correct.

CUSTOMER DATA HANDLING

There are a number of means of assigning customer loads - using and manipulating monthly billing values, using and manipulating total connected (badge) loads, or applying generic values based on customer type . How to handle and manipulate customer loads is beyond the scope of this document, however these values are often included in the GIS system or within a closely linked billing system or Customer Information System (CIS). So in that sense, in addition to the considering the traditional piping network, it is important to consider and handle the customer data when implementing a CAD-GIS to modeling interchange process.

At a minimum, customer data should include either a geographic or hydraulic location, and either a demand value or a link value that allows usage values to be extracted from a CIS or billing file.

EXPORTED DATA REQUIREMENTS

A network model is made up of many and varied data items. For example pipe size, connectivity, pressures, temperatures, and gas properties. A CAD-GIS system generally would only contain a small portion of the data used by a modeling application.

Although a CAD-GIS system does not contain all of the data needed to build a model, it can still be powerfully useful by providing the information required to create the foundation or framework to build a model from. At a minimum, a CAD-GIS system should be able to provide pipe location and connectivity, pipe size and length, and customer location and demand.

The exported data should contain information about the location of the pipe ends. This type of information can be in the form of geographic pipe location or coded pipe topology. Some modeling systems determine the pipe connectivity based on geographic pipe end proximity during the import process, in those cases actual connectivity is not required. Either the pipe location or topology is required for the modeling application to determine the connectivity of the resulting pipe network configuration.

Beyond the pipe configuration, certain attribute information is generally expected. For pipes, the pipe size and length values can usually be exported from the CAD-GIS. When being derived from CAD systems, the pipe size might not be explicitly defined, but might be implied based on the layer it is contained on or the line type used to displayed it, and the pipe lengths might be implied based on the geographic locations of the pipe ends and deflection points. From a GIS, the size and lengths can be either implied or derived similar to a CAD system, or they can be specifically defined in the attribute information.

Model node assignment is generally derived during the pipe import process. Node information such as pressure and load are not usually included in the CAD-GIS system. Generally no node information is required to be exported from the CAD-GIS.

Customer geographic location might be contained in a CAD system, however the load information would not normally be maintained there. Customer usage information is most commonly kept in a CIS or billing system. The CIS and CAD systems are not usually linked. On the other hand GIS implementations might contain both location, and demand or a link to the CIS. In either case, for it be truly useful, customer information should contain both geographic or hydraulic location and demand or usage information. The modeling application, should provide some means of interpreting this information and deriving loads from it that can be used to populate the model.

IMPORTING DATA INTO THE MODEL

Any modeling application which allows import of CAD-GIS data will likely provide some sort of interface to specify the format and details of the file to be imported. For example, assignment of layers to pipe sizes for a DXF file, or assignment of database fields to model parameters for a SHP file.

The import of line work is generally straight forward and widely supported. Some applications will allow specification of a “fuzzy” tolerance or similar parameter to help clean up “dirty” data. The fuzzy tolerance is a GIS term which indicates a zone around a node location. Any other nodes (pipe ends) found within the zone will be snapped to the original node. Node numbers or names are usually automatically assigned to the pipe ends during the import process. Pipe lengths may be calculated based on coordinate values, or assigned based on attribute values. Polylines and circular arcs may or may not be supported by the modeling application. Regardless of whether they are graphically supported, their true length should be considered when calculating pipe lengths. Pipe sizes may be handled by a layer assignment scheme or based on an assigned attribute value.

Node points are not necessarily required to be imported into the modeling system. Some modeling systems will automatically assign nodes at the pipe ends during the import process.

Regulator stations can provide a unique challenge to model creation. Many times a regulator station is depicted in the CAD-GIS system as a point object or entity. Accommodating a regulator in the network model, requires that the point be turned in to a hydraulic feature with a From Node and To Node - *and* that the flow direction of the regulator be known. Some modeling applications will provide a routine for automatically handling the import of regulators based on a certain syntax or data protocol. However in many cases, the placement of regulators may need to be manually performed.

System valves can present a similar issue as regulator stations, however their directionality does not generally need to be consider. Some modeling applications will allow system valves to be imported as a point feature, and do not require that they be converted to hydraulic elements in the model.

Not all modeling systems support the direct import of customer features, however all need to accommodate the import of the demands that they represent. Depending on the modeling application, the customers’ demand may be associated or assigned to a pipe, a node, a “meter” point, or some other load accumulating feature. Regardless of what the associated feature is called, the load ultimately ends up being assigned to a model node. This aside, the loads must somehow be imported and assigned to the appropriate associated feature. This association might be made by geographical proximity, for example to the nearest pipe or node. Or, by matching an attribute value, such as a pipe identification number or node name.

CHECKING AND CLEAN-UP OF MODEL DATA

After the data has been imported into the modeling application, it is important to check the resulting model. If the model was intended to be “to-scale”, one of the first checks is to check that the lengths and distances depicted in the model are representative of original map or drawing. This is an easy and important check, but is often overlooked.

Connectivity Errors:

CAD-GIS data often looks very nice, but is not topologically accurate. Topology refers to where and how the line work that represents the pipe segments are connected together. The topology of the source data basically defines the connectivity of the model. The most common issues with imported CAD-GIS model data involves connectivity. These issues can often be found using tracing or other automated tools in the modeling application, or by manual checking and review. Some common connectivity errors include...

Unsnapped Pipe Ends - This is a case where pipe ends are near each other but are not actually connected. Overshoots and undershoots are examples of this type of situation. See Figure 1 for an example.

Unbroken Intersections - This is a case where a “header” pipe is not broken at the connection of a lateral pipe. Often distribution systems are highly looped with many interconnections. The looping is often essential to the hydraulic performance of the system, and critical to obtaining accurate model results - and, can create real challenges to data checking. Intersection errors can greatly affect the quality of the results and need to be addressed and corrected. See Figure 2 for an example.

Overlapping Pipes - This situation has at least two cases. One where pipe ends that were intended to be connected, but are not connected because one of the other overshoot the other in direct alignment. The other situation is where a unconnected pipe segment overlays directly on top of a longer pipe segment. See Figure 3 and 4 for examples of this situations.

Dangling Pipes - This occurs when a single or multiple pipes are not connected to the remainder of the system. See Figure 5.

Zero Length Pipes - Occasionally during the digitizing process an operator will “double click” when entering a pipe segment, this results in a pipe segment with the same beginning and end location and of zero length. On import this results in a pipe with the same From and To node. This is generally considered an error in most modeling systems.

Extraneous Pipes - Sometimes extra line work is included on the same layer or in the same theme as the line work representing the piping system. This extraneous line work might depict detailed drawings, abandoned facilities, planned facilities, or mis-assigned line work (for example right- of-way lines on the pipe layer).

Extra Nodes - Depending on how the CAD-GIS database is created, many "extra" nodes may be present in the data. In this context an extra node would mean any node not necessary for the hydraulic analysis. This might include nodes at small service line taps - or nodes at a location where the pipe size and material is the same on both sides of the node, hydraulically the same, but the segments are considered different in the CAD-GIS system because they posses some different attribute value. For example they were installed at different times or under different work orders.

Complete Loops - Occasionally a complete loop will be used to depict a pipe segment, resulting in the same From and To node in the model. Mostly modeling systems would hydraulicly consider this an error. See Figure 6.

Attribute Errors:

In addition to connectivity errors a number of common issues can occur in the Attribute data. Some are describe here.

Pipe Size Errors - These errors occur when the pipe size and/or type contained in the GIS database are incorrect, or in a CAD system when a line is contained on the wrong layer. These types of errors are difficult to detect and usually require manual review to find.

Pipe Length - Assuming that the graphic scale is correct, these errors sometimes occur in data from a GIS where pipe lengths are assigned using an attribute value from the GIS database, as opposed to being calculated based on geographic locations. These errors are generally a result of data entry errors. These conditions can often be detected using automated features of the modeling system.

Null or Unknown Values - Often data contained in the CAD-GIS system is incomplete for example a pipe size or type might be set as "unknown" or there is missing data for older facilities. The model system is well equipped to handle these situations - for example, it is hard to calculate the pressure drop for a pipe of unknown size. These situations need to be addressed before the data can be useful.

Customer Related Errors:

Hydraulic data is handled more or less the same by most all applications used to model distribution systems - customer data is not. A few generically common issues are listed below, however because of these diverse modeling methods, finding the errors and handling the issues is not addressed.

Non-Gas User - Often billing data contains all "utility" customers, this could include water, wastewater, electric, and gas customers. When working with multiple service utilities, it is important to cull out or otherwise identify which customers do not use gas. For example, electric or water only customers should be removed from the database or otherwise handled in the load application within in the model.

Erroneous Billing Data - Similar to mapping or GIS most everyone believes that their billing data is perfect, and just as with those other types of data and systems, they are not. Often erroneous values are included in the database. When working with billing data, some means of scrutinizing the actual billed values needs to be implement by the user.

Location Errors - It is not only important to know the magnitude of a customer's load, but to have it applied at the correct hydraulic location. Distribution systems commonly have multiple mains in a street, or have multiple mains surrounding a customer location - for example maybe one or two different pressure mains in the street in front, a main in the alley behind, and a main in the side street on corner locations. It is important to ensure that the customer's load is applied to the correct main. The user needs to implement some means of detecting location errors.

Inactive Customers - While not an actual error there are nearly always cases where there are customers that have temporarily suspended gas usage. For example vacant homes, businesses, or factories. Depending on the type of analysis being performed, it may or may not be important to address these situations. For example, if an "existing system" analysis is being performed, no special handling may be needed. However if a some sort of planning analysis is being performed, it may be important to include these potential loads.

Linking To Attribute Data

Depending on the capabilities of the modeling system being used, it may be possible to link or attach non-hydraulic attribute data to features in model.

Most of the time a model of a distribution system is created to answer how or what-if questions. For example what-if a large load is added to the end of the system, or how is the best way to replace an aging system, or can a certain regulator station be retired, or can a main be taken out of service while a road or bridge is being worked on. These type of design questions often require that additional information be used to determine the most cost effective or practical solution. In these cases it might be desirable to know the test pressure, or installation date, or leak history of a pipe segment, or the highest or lowest risk ranking of various pipe segments. By linking non-hydraulic data to a model, the modeling system allows the user to quickly access all of the data that might be required to make these design decisions within the modeling system.

In order to link model and attribute data, the attribute data needs to be stored in a format that the modeling software can access and read. This can be done using the database portion of a shape file or through the use of other file formats such as text files, database files, or spreadsheet files. When working with a compatible file, the model needs to know which record in the attribute file is associated with which model feature. This can be done using a unique key or link data field or item. The link value needs to be maintained in both the model and attribute file. When attribute information is required for a feature in the model, the linked database is searched for a matching value, when a match is located, the associated attribute data is read from the attribute file.

Not all modeling systems support linking to an external attribute file, however if they do, it provides a powerful tool to the user, and provides an additional method of sharing GIS data with the modeling environment.

Exporting Data And Results

Once a model is created, and the what if and design scenarios analyzed and reviewed, it is sometimes desirable to share the results with a CAD system to create an attractive map or drawing, or to send the results to a GIS system for additional analysis. Most modeling systems allow the model data and results to be exported to the same interchange formats that were used to import the data that was used to create the model.

Exporting data to an outside application can provide a powerful extension to the traditional modeling system.

Maintenance & Updating

In most cases, a model created by data import will eventually need to be updated. There are several approaches to meet these requirements.

One approach is to periodically completely rebuild the hydraulic model by importing “fresh” data from the source. This method works well with fairly mature source data, that has been well vetted, and where errors found during previous imports have been identified and corrected.

Another method is to merge or append data that has changed or been added since the model was created or last updated. This method works well when date stamping and data segregation is implemented in the CAD-GIS, allowing it to be easily determined what data is new or revised. One issue with this method is that it is easy to add new facilities and make data value changes, however it is difficult to address removed or relocated facilitates or to handle partially modified features.

And as final alternative, a manual approach can be used to directly update the model data as changes are made to the actual piping system. Although this method causes some redundancy in effort, it sometimes proves to be the most practical method of maintaining a current hydraulic model. When using this method the CAD-GIS and model data tend to become more and more dissimilar over time. This is due to the manual entry aspect, due to differing levels of detail in the data depiction between the applications, and due to different interpretations of the source data by the different users. When using this maintenance approach, it is usually a good idea to periodically compare and reconcile the data maintained in the different applications, and eventually rebuild the hydraulic model from the reconciled source data.

Observations And Considerations

Following are few observations and issues worth considering when sharing or intending to share data between a CAD-GIS system and a network modeling application...

Applicability - Although the primary focus of this document has been on working with distribution style piping systems, the basic issues and points addressed apply to most any style piping system - topology and connectivity need to be correct, a coordinate system needs to be selected, attribute data needs to be accurate, devices need to be handled, and supplies or demands to or from the system need to considered. All these issues are applicable to any distribution, transmission, gathering, or plant piping system.

Database Design - No organization wants to hear that the CAD or GIS data that they have spent thousands or millions of dollars to build isn’t all that special, but many times that is the case. The graphic data nearly always looks good, but the topology is poor, the data accuracy is poor, the coordinate system is not practical, or the database doesn’t contain the data items required to support critical applications like network modeling. Almost every poorly designed or implemented CAD-GIS system can be corrected, but at a cost. It is important that the considerations to support applications like hydraulic modeling be addressed early in the CAD-GIS database and system design.

Data Quality - Do not rely on the data imported from CAD-GIS systems to be “perfect” or even all that good. Always check the data before trying to use it. And when errors are found, correct them in the source, so that the next time the data is imported it does not require re-editing. By using this method, over time, more confidence can be placed in the imported data and less effort will need to be expended to check and correct the data.

Pipe Size and Type - Depending on the modeling system being used, it may require the input of actual inside diameter values, or it may allow the input of codes representing size and material type values. In either case, the required value needs to be imported into the modeling system. Often the pipe size in the GIS database is stored as a nominal value, not an actual inside diameter value - this presents challenges to modeling systems that require the import of actual inside diameters. On the other hand, the material type in a GIS database is often maintained as a normalized code, for example 1 = polyethylene plastic, 2 = steel - this likewise presents a challenge to modeling systems that use a size/type code. Often these challenges require some preprocessing of the data before the data is imported into the modeling system.

Hydraulic Length - When importing data from a two-dimensional CAD file, the hydraulic length is equal to the horizontal graphic length of the line work found in the original file. When importing data from a three-dimensional CAD file, depending on the modeling system's capability, the hydraulic length can more closely reflect the "true" hydraulic length by using the full X-Y-Z coordinate triples found in the original file.

When importing data from a GIS system the hydraulic length can be assigned to either the 2D or 3D graphic length, or to an attribute value. Some operators maintain separate lengths in the GIS database, for example the "as-laid" or "inspected" or "measured" lengths are kept as data items. These values are almost always different from the graphic length. A decision needs to be made as to which length value is appropriate to use in the modeling system.

The precision of the graphic length value is not completely critical to developing a useful hydraulic model, variations on the order of a meter or of a few feet will make little difference in most modeling results. However it is important, when using imported data, to remember to check the graphic scale of imported data to at least ensure that it is of the correct magnitude. Always check the scale before spending resources editing or using bad data.

Normalized Databases - One method of database management and definition uses something referred to as normalization. A normalized database is made up of a variety of separate tables, sometimes called lookup tables, that are linked to each other through codes and key fields. Often the values maintained in the "database" do not represent the actual value of the item, but might be represented by a numeric code. For example, to represent polyethylene pipe, the material might be represented by the number one, instead of the actual value of "polyethylene". Because there is not a single table representing the data, this type of database configuration causes some challenges to importing data, and the data is not always represented by values recognizable by the modeling system. Pre-processing is nearly always required before attempting to import data from these style of databases .

Date Stamping - Keeping track of when a change is made in the CAD-GIS system allows for effective updates to the network model. In GIS systems it is good practice to include a data field for maintaining entry or revision date. In CAD systems separate layers can be used to identify pipes installed during certain periods.

Data Segregation - It is nearly always easier to combine data when it is required, then it is to separate data when it is not required. More segregation during the design and implementation of a CAD-GIS system is a good thing. Storage is inexpensive, and most all CAD-GIS and modeling systems provide some mechanism to automatically combine data, so consider maintaining the CAD-GIS data in many logically separate layers or themes.

Unsupported Data - Data imported from a CAD system will generally be limited to pipe size and length. There is significantly more data required to develop a hydraulic piping model. For example, pipe flow equation, efficiency, gas properties, and gas temperature. Similarly a GIS system would not generally maintain all of the data items required by a hydraulic model. In order to create a functional model, these missing data need to be accommodated. Depending on the modeling system, automated tools may be available to set these values to default values or en-masse.

Elevation - Elevation values are not always maintained in the CAD-GIS systems. In most circumstances, elevation differences in distribution systems are not critical to hydraulic analysis. For "pounds" systems, use of an average value is generally acceptable. However for low pressure "inches" or "ounces" systems, it is important to capture significant elevation changes. In low pressure systems the change in elevation can cause a noticeable change in the absolute pressure due to the associated change in atmospheric pressure. The atmospheric pressure is a function of elevation, and absolute pressure difference is used by all pipe flow equations.

Model Verification And Calibration - Pipe diameters are difficult to check, however it is important that they are correct. Of the values that are used in a hydraulic analysis, diameter is one of the most sensitive. Because of how the diameter is used in the pipe flow equations, small changes in diameter, cause large changes in the hydraulic results. Problems with pipe sizes are found by luck, by careful review of the hydraulic results, and by comparison of model results to actual field values or during "calibration" of the model. Where practical it is always good practice to verify or calibrate the model using actual field measured data. These procedures will help find any significant diameter or other configuration or data errors.

Documentation - The successful integration of a CAD-GIS and modeling system often hinges not on the technology, but on good and thorough documentation. The supporting documentation should address issues such as: minimum required database design and definition, drawing layout and layer scheme, coordinate and projection system to be used, dimensional units to be used, and a step-by-step guide or checklist on the export, import, and data checking procedures.

CONCLUSION

Although CAD-GIS and network modeling systems have not evolved into a single technology, there is a great need to share data between the systems. Consideration of the issues highlighted here, will help ensure that the data integration between the technologies is successful.

REFERENCES

None

ACKNOWLEDGMENTS

None

AUTHOR BIOGRAPHY

Brad Bean is a partner in a firm that goes by his name - Bradley B Bean, PE (B3PE). The firm provides engineering and software services to the natural gas industry. He has a Bachelors of Science degree in Mechanical Engineering from Colorado State University and is a registered engineer in several states. After college he spent a few years with Stone and Webster Engineering performing pipe hanger analysis for nuclear power plants. He started in the gas business in 1982, spending 10 years with the City of Colorado Springs Utilities (CSU). During his time with CSU he was responsible for implementing the gas utility’s first network modeling application, first AutoCad mapping and construction drawing applications, and participated in implementing their ESRI based GIS system. When he left CSU in 1992, he founded B3PE and has directly participated in the firm’s operation since that time. He has personally worked on the development of their network modeling and other design and engineering applications, and still works as an active engineer on the firm’s various modeling and design projects.

FIGURES

Figure 1 - Unsnapped Pipe Ends

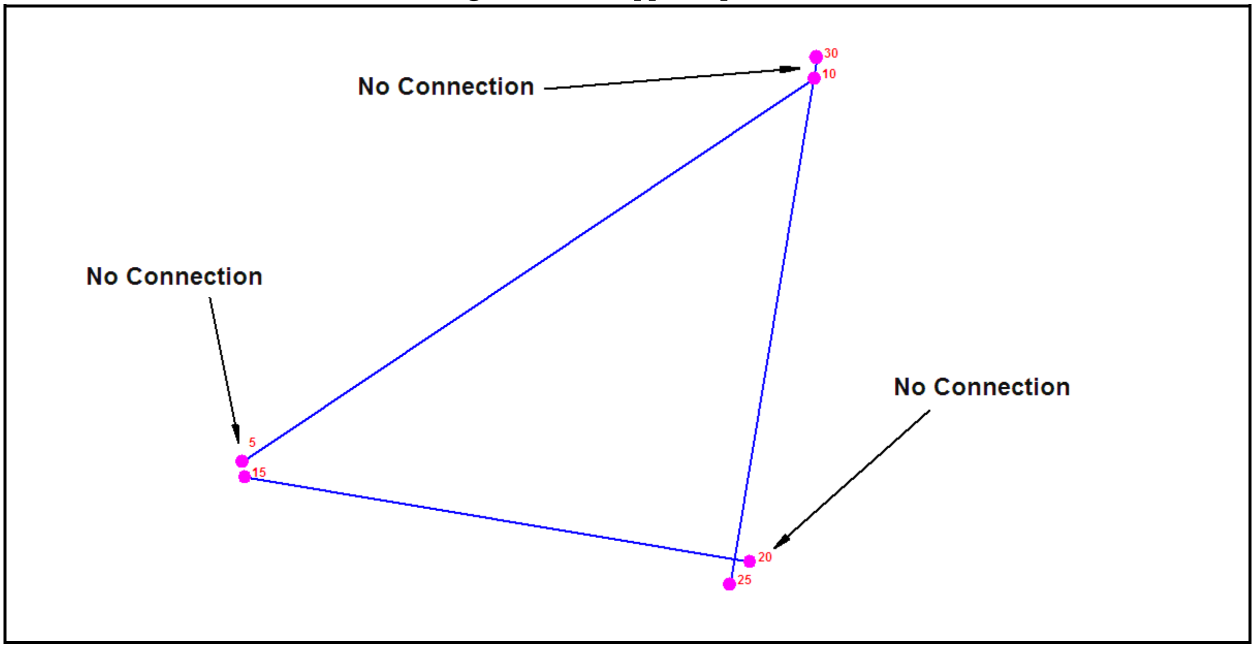

Figure 2 - Unbroken Intersection

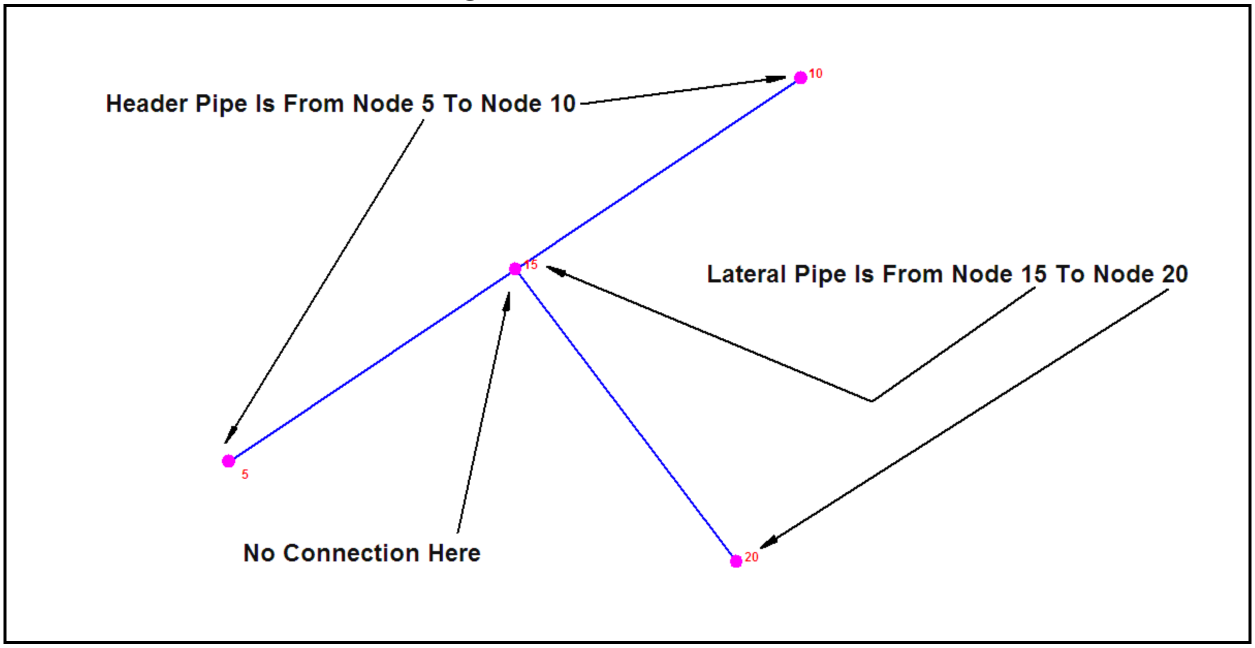

Figure 3 - Overshoot Example 1

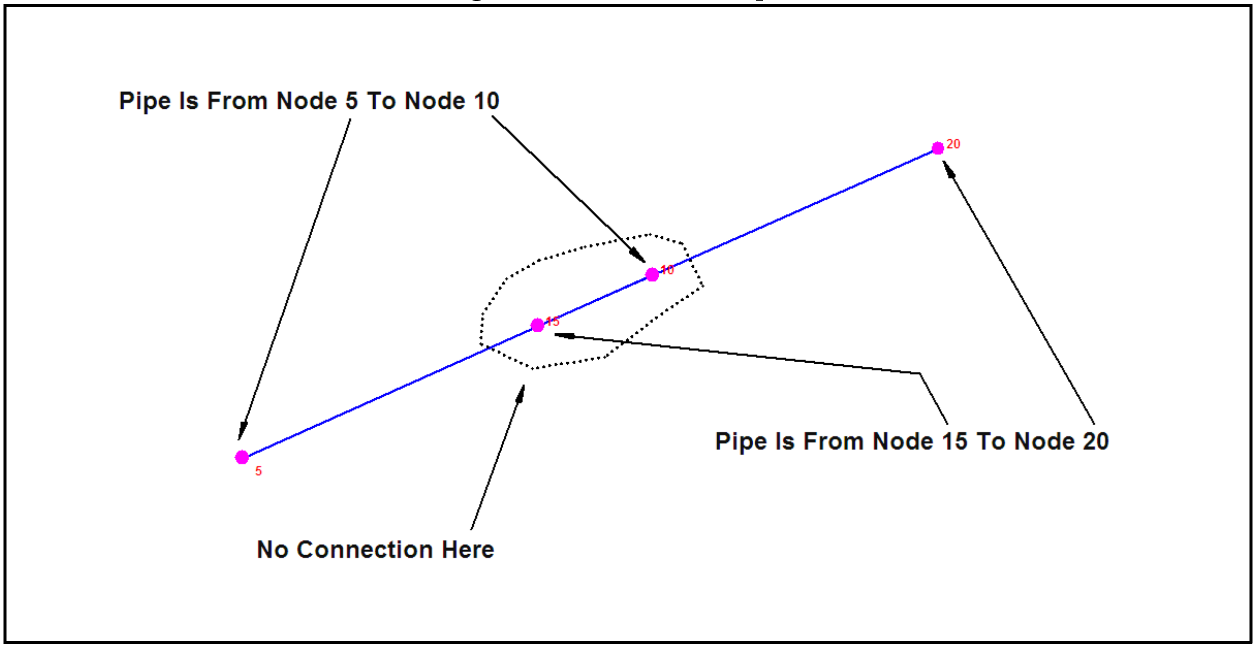

Figure 4 - Overshoot Example 2

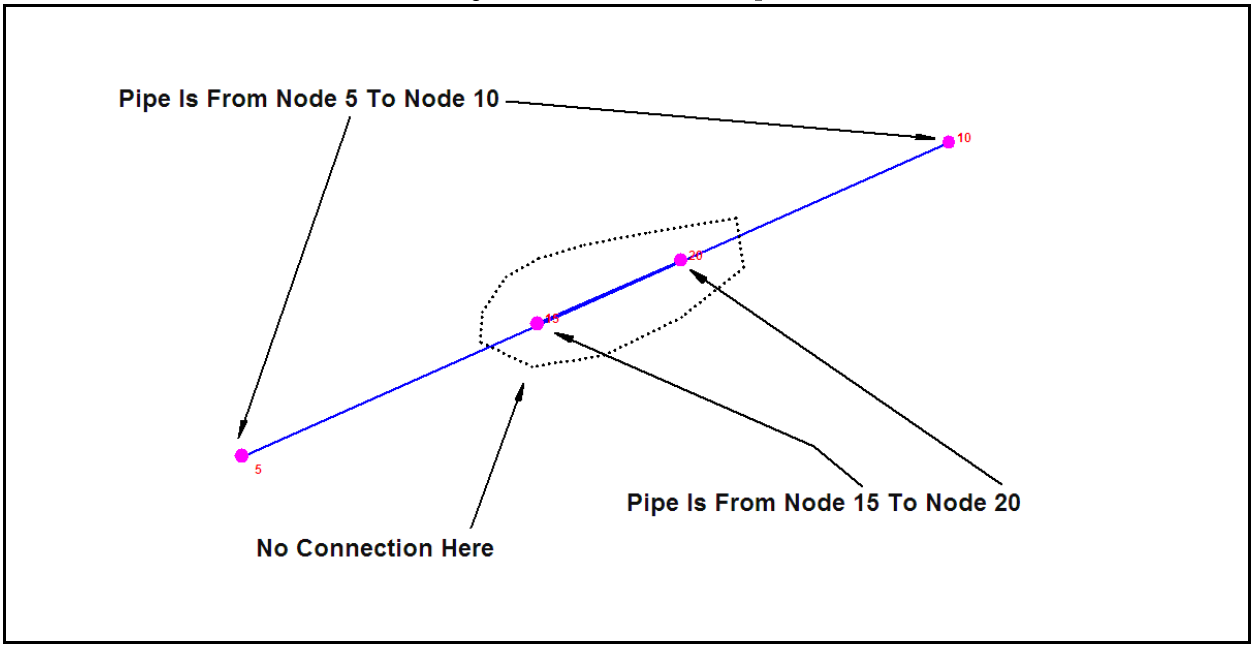

Figure 5 - Dangling Pipe

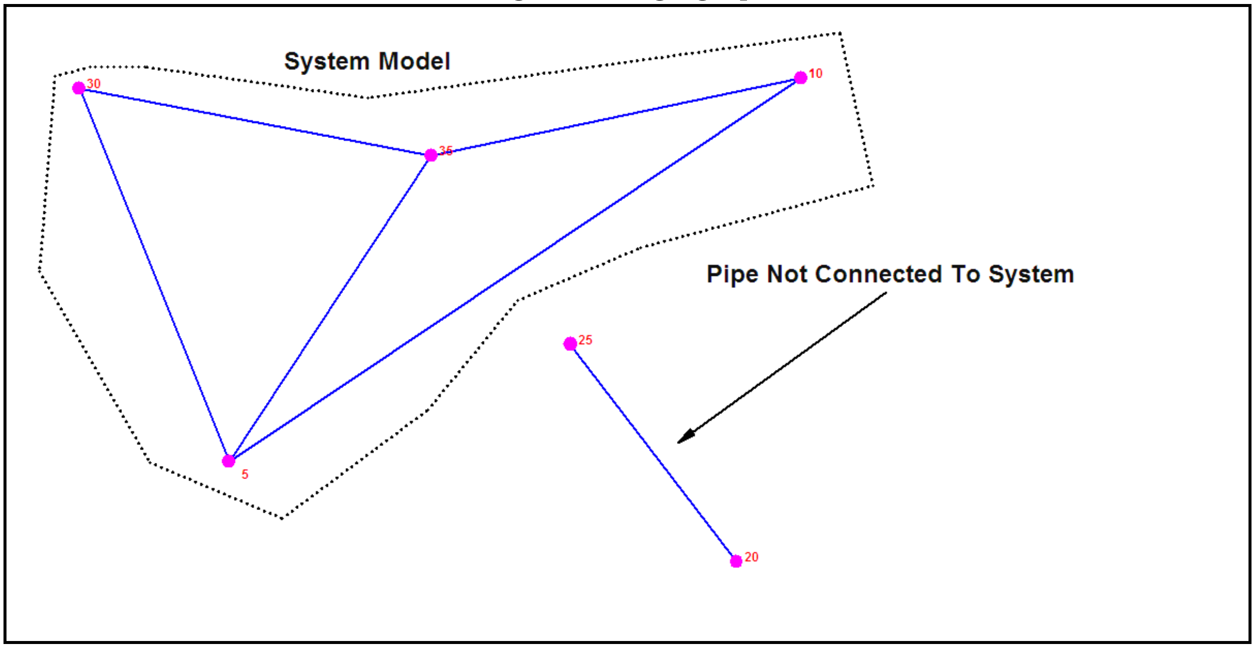

Figure 6 - Complete Loop