EXPLOITATION OF DIVERSE GIS DATA MODELS FOR ROUTING AND OTHER PLANNING PURPOSES

Zlatko ZMIJAREVIĆ
HERA[1] - Croatia
zlatko.zmijarevic@hera.hr

Minea SKOK
FER[2] - Croatia
minea.skok@fer.hr

Hrvoje KEKO
FER[2] – Croatia
hrvoje.keko@fer.hr

Davor ŠKRLEC
FER[2] – Croatia
davor.skrlec@fer.hr

Slavko KRAJCAR
FER[2] - Croatia
slavko.krajcar@fer.hr

ABSTRACT

Today, electric distribution planning methods rely on precise digital data readily available through enterprise wide information systems. Such data does not encompass just the electric power system, but also covers selected georeferenced data related to supply areas. This paper deals with using spatial data for long-term planning of medium voltage networks related, in particular with feeder costs and routing. In fact, the Authors have been developing an optimization method for planning open-loop and link networks based on evolutionary algorithms. For exploiting geographical data a module has been produced employing Geographic Information Systems (GIS). Although the primary role of the Module is to tie the optimization model with real data (actual objects) and provide feeder routing, the Module provides other useful features like adaptive use of data, advanced modelling of the supply area for routing, flexible routing schemes, etc. Within this paper, the focus is placed on different routing schemes based on diverse data models and data sources built upon the planning method. Different cost models for constructing feeders are presented together with explanations and guidelines for utilizing those models and associated spatial data. Routing models presented here are not limited to the underlying planning method but are generally applicable to any planning method or automated routing algorithm. The described tools herein, together with procedures and GIS modelling, may provide a starting point for planners to devise and use their own geographical models.

INTRODUCTION

The Authors have been involved in devising distribution planning methods and associated computer software for use in the Croatian electricity company (Hrvatska elektroprivreda d.d.) since the late 1980-ties. The primary objective was to produce a tool for long-term planning of medium voltage networks in urban areas using link and loop networks as preferred network layouts. The name CADDiN was adopted for the method as a clear indication of a computer supported method (CADDiN is an acronym for Computer Aided Design of Distribution Networks).

The CADDiN method is built upon a medium-voltage (MV) network layout optimization module that minimizes overall expenditures. Given the locations of load points and their forecasted power demand, the optimization module will provide a best possible MV link or open-loop layout according to imposed technical constraints.

It should be emphasized, that the optimization algorithm is concerned with sequences (or cycles) of load points which determines the overall network (open-loop) layout. This is often considered as “routing”. However, in this paper the word “routing” will be interchanged with “determining the actual route between two load points”. To elaborate further, the optimization relies on a proprietary mathematical network for finding the best possible electric distribution network layout. In this mathematical network each edge between any two load/supply points has an economically defined cost and certain electric properties. Actually, the optimization uses a complete graph (every two vertices in the network are connected by an edge), so the planner does not need to define probable options (new feeder candidate). This aspect must be emphasized since many other planning methods depend on options defined in advance which may lead to suboptimal solutions (a planner may omit an opportune option). The optimization algorithm constructs candidate layouts and searches for the one with the lowest cost. Single-stage distribution planning is possible by using a cost between two load points representing fixed and variable expenditures for such a link during the planning period. The method is described in [1,2,3,4].

CADDiN uses GIS data through stand-alone modules (interfaces). Although GIS is a most suitable environment for planning purposes due to a number of reasons (graphical interface, data access, spatial selection/analysis, topological queries, etc.) it is not a practical platform for difficult combinatorial optimization – programming interfaces of GIS tools are not designed for such extensive computation. In general, direct utilization of GIS data by automated planning procedures may be challenging depending on the underlying planning method (i.e. those using fuzzy numbers or probability data to model uncertainties). Furthermore, this separation allows the use of any arbitrary software and source of data.

CADDiN GIS modules have been evolving from initial automated data preparation procedures in AutoCAD 10. Modules for use with older GIS software *System 9* and *ArcView* have been created. The most recent module named *MapModul* is based on the desktop GIS software *Autodesk*

1 HERA – Croatian Energy Regulatory Agency
2 FER – Faculty of Electrical Engineering and Computing, University of Zagreb

Map and a *Smallworld GIS* module is under construction. The most important functionalities of these modules are:

1. specifying elements of the planning area to be used in the planning process (load/supply points, existing feeders, available corridors);
2. constructing of the proprietary network for optimization (as indicated earlier) with cost and electrical data;
3. creating files necessary for the optimization module;
4. displaying results from the optimization module.

The *MapModul* has additional features that include adaptive use of data, advanced modelling of the supply area for routing, flexible routing schemes, etc. Although, some details on the *MapModul* are presented in this paper, more information may be found in [5].

CADDIN MAPMODUL

Autodesk Map was recognized as a promising platform since it combines GIS functionalities with the world's most popular CAD software *AutoCAD*. The primary incentive for selecting *Autodesk Map* had been the ability to handle at the same time true GIS data and legacy data (stored in CAD drawings). As a desktop GIS, *Autodesk Map* is designed for using different sources of data (acquired by any means available) and may interface with enterprise GIS software. Built-in interfaces as well as various import/export routines for using GIS data in other formats allow *Autodesk Map*, and consequently the planning method, to use a wide variety of data sources. An accompanying advantage of using a desktop GIS is that it's independent from enterprise-wide information systems and therefore appealing for planners who create and devise numerous sets of potential problems and solutions without the danger of changing the original data. However, downsides may appear if repetitive procedures are needed in order to retrieve up-to-date data.
As mentioned earlier, the *MapModul* determines the best route for every two load/supply points. Existing feeders that satisfy given planning criteria are considered best routes since they are already constructed. The optimization algorithm will ascertain if a existing feeder should be abandoned. If no existing feeder connects two points, a new feeder is proposed for the optimization algorithm by an automated procedure in *MapModul*.
In CADDiN, routing feeders over the terrain is based on a network of allowable corridors, where corridors are line entities on which feeders may be constructed. Corridors may be tied to any available line entity - street corridors or centerlines, cadastral boundaries, etc. Routing is constrained to the network of allowable corridors because the planning method is adapted to urban areas where routing is not normally allowed over private or municipal property. This last argument is increasingly significant due to environmental and property issues in developed countries. Avoiding natural and artificial obstacles is a simply a matter of inserting, deleting or modifying GIS line entities. On the other hand, routing feeders in rural areas or routing transmission lines is not that restricted, but is challenging in its own way [6,7]. Regardless of the model, main reasons for employing more sophisticated methods for determining feeder routes in optimal distribution planning are: 1) to provide long-term planning methods with more accurate data (costs and electrical data that primarily depends on feeder length) and 2) provide mechanisms for avoiding spatial obstacles. Just to be comprehensive, simple models are elaborated in [8] (simple grids), [9] (geometric modelling). In earlier versions of CADDiN, feeder costs were based on Euclidean or max norm distances between load points located in a grid.
All in all, using specific cost models, as shown latter on, routing can take into consideration cadastral data, terrain data (vegetation coverage, soil type, elevation, etc.) and other data.

EXPANDING ON AUTODESK MAP

Since *Autodesk Map* evolved from a CAD tool, it differs somewhat from other desktop GIS tools (lightweight GIS software). In order to preserve compatibility with *AutoCAD*, *Autodesk Map* stores and handles geographic data in *AutoCAD* "drawings". But instead of directly using data in an active drawing, drawings with data are "attached" to an active drawing. These "attached" or source drawings provide a repository of geographical data. By using so-called "queries" *Autodesk Map* extracts solely entities of interest into the active drawing, using GIS functionalities (analyses) – see Figure 1. These "queries" may use various criteria for selecting entities:

- spatial criteria
- data attached to entities,
- CAD properties (i.e. name of layer, colour),

Criteria may be combined by logical operators to provide more complex queries which users may use to make very precise selection of objects that comprise to planning problem. Additionally, *Autodesk Map* queries may include SQL queries for more intricate data extraction.
All this is similar to conventional GIS software, except *Autodesk Map* can handle different types of entities despite different data being attached to entities. Finally, the "query" mechanism is equipped with a transformation mechanism that can change certain CAD properties of entities. A nice example is using *AutoCAD* blocks interchangeably with points, allowing the creation of an integral geographic database based on several previously mismatched area databases with legacy data.

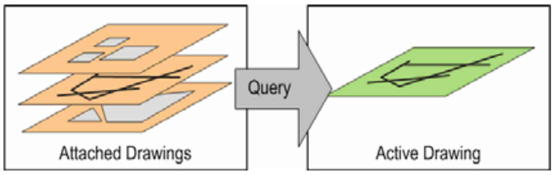

Figure 1. Extraction of data from attached drawings

The *Autodesk Map* concept of "attached drawings" and "queries" is helpful in planning purposes, since users may handle different variants of the planning problem (i.e. perform decomposition). Moreover, "queries" can be stored, reused and shared between drawings or planning variants.
In addition to these out-of-the-box data related features, *Autodesk Map* is equipped with an important mechanism that significantly increases flexibility in handling data. The mechanism is called the "expression evaluator" and allows the user to compute data based on a given GIS entity and a valid expression. Figure 2 illustrates the computation of a

simple expression on a line entity.
"Expressions" are formulae for computing values using:

- entity properties (*AutoCAD* properties or geometry information like length or area)
- data attached to entities (data stored in native or external databases)
- legacy data (data stored in AutoCAD blocks)
- user variables, etc.

"Expressions" can be extended by arithmetic functions, string functions, logic operators, conversion functions, etc. "Expressions" are based on *AutoLISP*, a built-in interpreter language for operating *AutoCAD*.

Figure 2. Extraction of data from attached drawings

The CADDiN *MapModul* extensively uses "expressions" and "queries" in data preparation for the CADDiN optimization module. Instead of enforcing its own data model, during the data preparation the *MapModul* transforms existing (source) data using "expressions" into data used in the optimization. This allows the *MapModul* to use any existing data (the data requirements of the CADDiN method are undemanding in contrast to possibly very complex data models used in the utility). "Queries" are used for preparing an active drawing representing the planning problem at hand.
Some of the procedures used for data transformations in *MapModul* have been exposed as accompanying tools. These tools, together with and built-in data capabilities of *Autodesk Map*, provide a very flexible environment for handling data without programming. Especially useful is the *MapModul DataFill* tool for evaluating an "expression" on a set of object and writing the results back to the objects. This tool has been especially useful for handling and combining partial sources of inconsistent legacy data. An illustrative example from real life is where 10kV and 20kV medium voltage cables in a DXF file differ only in colour.

TERRAIN COST MODELS

After explaining the routing concept of CADDiN and certain mechanisms of *Autodesk Map*, the actual routing model can be unveiled. As a reminder, CADDiN uses a complete graph of feeders connecting pairs of load points for optimization, while the *MapModule* prepares one based on an artificial network of GIS line entities. The artificial network of allowable corridors is constructed on any data source available to the planner (street centrelines, existing feeder routes, manually inserted corridors, etc.).
This emphasis on "*any available source*" is in fact a necessity from real situations. The complexity and size of the utility's GIS (the actual distribution system with each and every component) makes the implementation, development and maintenance of the utility's GIS cumbersome and consequently expensive. Consequently, it is not surprising that the utility's GIS data envelopes all (or as much as possible) of the distribution system, minimal data concerning the real world (possibly the road network or cadastral information), but scarcely any other data (usually small amounts of data associated to some other business processes within the utility like assets, organizational elements, etc.). Unfortunately, planning activities rely on information concerning the area containing the distribution grid (like urban planning information, objects of other systems like gas networks, demographic data, etc). This kind of data is hard to obtain and of small value to other business activities in the utility. On the bright side, implementation and development of a utility's GIS typically involves assimilation of data under the authority of other companies or institutions (i.e. road networks). Within the assimilated data, there might be some valuable for the planner. For that reason, most planners must use what is available or create their own datasets.
Routing on the MapModul's network of allowable corridors is simply a matter of employing a Shortest Path algorithm (see [10] for definitions and algorithms). Fortunately, most GIS software is equipped with some sort of shortest path algorithm for basic topological analyses. The prevailing advantage of *Autodesk Map's* built-in shortest path algorithm is the utilization of "expressions" for defining costs attached to edges (and vertexes). Most GIS software uses length or a user-selected numerical field as the cost for the shortest path algorithm. Using "expressions" in Autodesk Map is far more flexible since "expressions" may use functions and different data. To achieve the same in some other environment would call for data manipulation (to fill a database field with the appropriate cost beforehand) or some sort of programming.
Using appropriate "queries" CADDiN *MapModul* will first create a set of line entities forming a network, after which costs will be calculated for each edge in that network based on user defined "expressions". This way, by customizing "expressions", the planner may tune the cost model for routing to his/her data. However, extravagant models are not necessary for planning because construction costs fluctuate considerably (specific project cost, uncertainties, etc.). It must be mentioned that regardless of these fluctuations, planning still attempts to make more precise estimates. In densely populated areas, alternative routes between two points may drastically differ in cost.
Simple, yet effective cost models are easy to perceive and implement. First of all, a planner may choose between complete costs for each edge (corridor) or use *per unit* cost and multiply it with length (i.e. per meter cost). The first option involves using costs that do not depend on length, which is not that common. The second option is more common. In both options, terrain specifics may be introduced.
A basic terrain model or estimating feeder cost may envelop a base cost and modifiers related to 1) soil type, 2) surface restoration (road surface type or simply road type) and 3) vegetation (for overhead lines). Administrative fees may be considerable for new routes, so a modifier is viable that will modify edge costs for existing feeders (in Croatia administrative fees and right of way settlements are around 25% of the overall cost). Slope may also be introduced as a modifier, although 3D terrain models and data may be quite expensive and therefore unavailable to the planner. These modifiers may be easily defined as factors and multiplied with the base cost. This approach may be perceived as **multiplicative**.
An **additive** model of costs may be a sum of individual

costs (trenching costs, surface restoration costs), each calculated separately. The difference between multiplicative and additive approaches is based on the difference between using overall and per unit costs on an edge.

The *MapModul DataFill* tool mentioned earlier in combination with spatial queries and polygon overlays (GIS analyses like line-in-polygon) is quite handy for creating appropriate data (i.e. providing cost factors for corridors in areas with different soil-type). Since "expressions" work on individual GIS entities, the planner may modify data attached to a specific object in order to influence routing. A real-life example is setting extremely high costs for old city streets. The ability of the *MapModul* to easily change cost models without programming is most beneficial in creating and testing new planning models in comparison to classic GIS approach, where preparing appropriate data involves manipulations with layers and joins. While the basic idea is the same, achieving the final result is more cumbersome.

For instance, a planner may use a uniform per meter cost for constructing feeders. These costs may be easily modified by a zoning factor awarded to a district, i.e. higher costs for city centre. Using simple polygon selection the planner may easily set the zoning factor.

Another example is creating a surface restoration factor based on an enumeration type of a road. For instance, roads with asphalt cover have 20% higher cost, roads with stone blocks have 100% higher costs, etc. This example requires some initial effort but afterwards it is easily reusable in planning purposes.

Finally, to conclude the CADDiN planning paradigm, single stage planning in CADDiN is possible by multiplying estimated costs of feeder construction with an economic factor. To be more precise, the original cost is split into uniform annual series and increased by a maintenance factor (expressed as a percentage of the investment cost), then translated to present value using a discount rate. The applied discounting techniques are further elaborated in [11].

CONCLUSION

Given the possibilities of Geographic Information Systems, great expectations have been placed upon GIS applications in electric utilities, especially planning. Nevertheless, GIS data for planning may not be readily available. This paper showed how limited geographical data may be used in combination with a long-term planning tool. The planning tool is based on separate modules – a core optimization module and different GIS modules for utilizing GIS data. Based on the features and flexibility of the GIS module, different cost model for routing over terrain have been explored and consequently used in practice.

All things considered, the CADDiN planning method and the *MapModul* proved to be useful for planning in situations where data is incomplete, insufficient, etc. In addition, the routing mechanism is adaptable allowing the planner to produce custom routing models based on any geographical data.

REFERENCES

[1] Z. Zmijarević et al., 2005, "A Comprehensive Methodology For Long-Term Planning Of Distribution Networks With Intrinsic Contingency Support", *CIRED 18th International Conference and Exhibition on Electricity Distribution*, Torino, Italy, Published on CD-ROM.

[2] M. Filipec, D. Škrlec and S. Krajcar, 2000, "Genetic Algorithm Approach for Multiple Depot Capacitated Vehicle Routing Problem With Heuristic Improvements", *Int. Jour. Of Modeling and Simulation,* Vol. 20 No. 4, 320-328.

[3] M. Skok, D. Škrlec, S. Krajcar, 2001, "Comparison of two Genetic Algorithms for Link Distribution System Design", *World Multiconference on Systems, Cybernetics and Informatics SCI 2001*, Orlando, USA, 23-29.

[4] Z.Zmijarević, M. Skok, D. Škrlec, S. Krajcar, 2001, "Application Of A Genetic Algorithm Based Distribution Planning Method For Creating Diverse Link Layouts", *Proc. of the 4th IASTED International Conference on Power And Energy Systems (EuroPES2004)*, Rhodes, Greece, 406-411.

[5] Z.Zmijarevic, H.Keko, I.Muzic, M.Skok, D.Skrlec, S.Krajcar,2006, "Data-model independent GIS module with flexible and robust routing for optimal long-term planning", *Proceedings of CIDEL 2006,* 27-29. Nov. 2006, Buenos Aires, Argentina.

[6] N.A.West, B.Dwolatzky, A.S.Meyer, "Terrain-Based Routing of Distribution Cables", *IEEE Computer Applications in Power*, Vol.10 No.1, 1997, 42-46.

[7] C.Monteiro, I.J.Ramirez-Rosado, V.Miranda, P.J.Zorzano-Santamaria, E. Garcia-Garrido, L.A.Fernandez-Jimenez "GIS Spatial Analysis Applied to Electric Line Routing Optimization", *IEEE Transactions on Power Delivery*, Vol.20 No.2, 2005, 934-942.

[8] D.L.Crawford, S.B.Holt, "A Mathematical Optimization Technique For Locating And Sizing Of Distribution Substations, And Deriving Their Optimal Service Areas", *IEEE Transactions on Power Apparatus and Systems*, Vol 94. No. 2, 1975, 230-235.

[9] T.H.Fawzi, K.F.Ali, S.M.El-Sobki, "Routing Optimization Of Primary Rural Distribution Feeders", *IEEE Transactions on Power Apparatus and Systems*, Vol.101 No.5, 1982, 1129-1133.

[10] T.H.Cormen, C.E.Leiserson, R.L.Rivest 1997, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, USA.

[11] L.T.Blank, A.J.Tarquin 1998, *Engineering Economy*, 4th ed., McGraw-Hill , Boston, USA.