Using schematic mapping for synthetic networks

David Fairbairn
Newcastle University
Newcastle upon Tyne, UK
david.fairbairn@newcastle.ac.uk

Qingyuan Ji
Newcastle University
Newcastle upon Tyne, UK
Q.Ji2@newcastle.ac.uk

***Abstract*—This research addresses the development and application of a standard heuristic algorithm to generate synthetic networks for representing underground infrastructure networks at an urban scale. Such networks are required to model the unknown locations and trajectories of the local electricity supply grid serving a city-wide area. The graphical output from such activity is a pseudo-schematic map, which can be applied in visual examination for validation purposes, and as a means of assessing network dependencies and loading.**

Keywords—synthetic networks, 'big data' networks, electricity networks, heuristic network algorithm

I. INTRODUCTION

The modelling of infrastructure assets as network nodes and edges (both physical, such as pipes and cables, and intangible, such as wireless telecommunication) is long-standing, logical, and indeed necessary for efficient planning, management, maintenance and expansion. Once topological network connectivity is established, it is possible to perform more complex tasks on the system, such as infrastructure network vulnerability analysis [1] and cascading failure simulation [2]. It is also possible to use network theory to model the interdependency of different infrastructure networks, where the state and operation of assets within one type of infrastructure network (e.g. water supply) may depend on those within another network (e.g. electricity).

Underground infrastructure management suffers due to historical inattention to recording assets, leading to incomplete network maps. Both commercial confidentiality and awareness of possible malicious interference with assets, may also limit any availability of network maps for public perusal or third-party management.

II. SYNTHETIC NETWORKS

A. *Rule-based water network derivation*

Exemplifying the addressing of the first problem of incomplete maps, Northumbrian Water Limited (NWL), Durham, England, commissioned the geospatial company 1Spatial to derive a true-location network plan for their water supply and waste-water disposal infrastructure, virtually all underground, but with very limited current records [3]. The solution was to take the known locations of existing surface assets (inspection covers, pumping stations, supplied buildings, sources of waste-water) and interpolate the route of the underground pipe-work. Much processing was based on *guidelines* (e.g. no pipes run underneath a building; flow direction is always downhill; properties >1 km from a sewer are not connected to the main network) and *expert system derived rules* (e.g. sewer pipe Runs down the back alley IF property=terrace AND age 1900-1930) which required significant additional information (e.g. building age).

B. *Developing graphical algorithms*

In synthetic network generation, information may be more limited. The heuristic used in this research, for the inferring of the electricity supply network, assumes the only known information as the layout of buildings, accurate location of some assets (notably distribution sub-stations), and routes (notably paved roads) used to link them. Further, the total length of the infrastructure networks should be kept as short as possible, and they are always buried within a road polygon, either on the centre-line or at the edge. Moreover, we assume spatially-close buildings in Euclidean space should be served by the same infrastructure assets, and that each building has only one connection to the network. Finally, when generating the spatial layout of networks, a geometry check is undertaken to make sure the actual building footprints (polygons) access the network in a valid way.

III. ALGORITHM OUTPUTS FOR NETWORK GENERATION

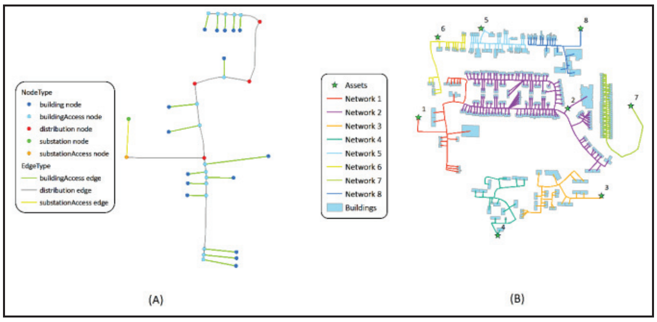

Figure 1: (A) defining terminology of node and edge types (B) results of applying the algorithm to a local district with several separate electricity substations (Assets)

The results of the networks generated are stylised and equivalent to a schematic map: the algorithm assumes straight line connection between buildings and neighbouring distribution cables in the nearest road centre-line, although there are rules forbidding cables located underneath buildings either *en route*, or through adjoining properties. In reality, the actual distribution feeders normally follow one side of the road (i.e. asymmetrical within the road polygon) (Fig 2). Thus, a road segment measured as the nearest to a building based on distance from the road centre line, may not necessarily be the closest if distance to a nearside of the road is considered.

For assessment of the performance of the algorithm in this exercise, some real-world data was released by the local electricity supplier solely for this purpose. Concerns addressed include realism of the network routes, in particular the nature of the connection between the building and the main distribution cable, and the viability of the loading on each asset (i.e. number of buildings attached) (Figs 3, 4).

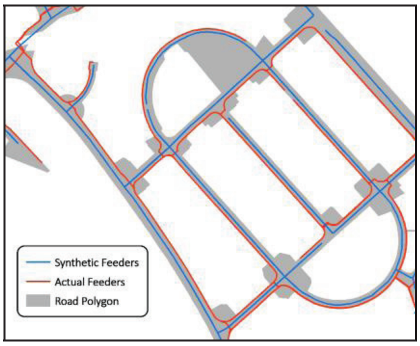

Figure 2: algorithm assumption is that the distribution cables adhere to the road centre-line, but in reality they are more efficiently installed at the edges of the road

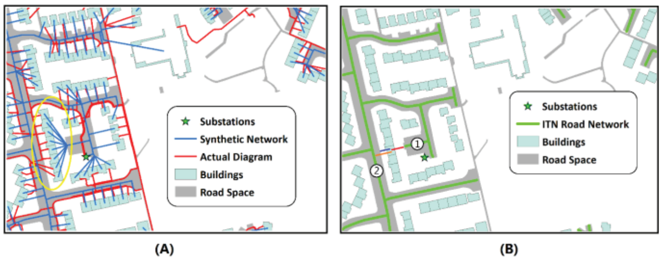

Figure 3: (A) mismatch between synthetic network and actual diagram caused by different definition of distance between building and road (B) [blue line connecting to road-edge location of network cable at 2 is shorter than red line connecting to network at 1, so layout in (A) is suboptimal

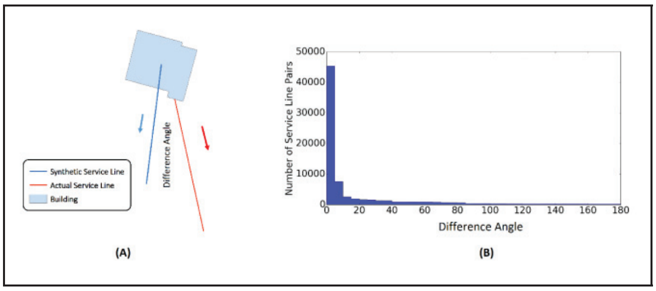

Figure 4: assessing accuracy of angular orientation of connection to building

The validation has, therefore, been quantitative in some respects, but is also partly visual in nature – acknowledging that most interaction with schematic diagrams is through human agency is important.

IV. Applying the Algorithm

To ensure its transferability and scalability, notably in handling big data sets at large city scales, a schematic network was derived for the whole of the city of Newcastle upon Tyne, England (636 sub-networks created, 5 hours 12 minutes algorithm run time), and also for the large urban area of Greater London. For the latter, 2,239,213 buildings connected to the electricity network were identified from the Ordnance Survey database, along with 16,839 electricity substations feeding the distribution network. On a desktop workstation, the algorithm took 156 hours 27 minutes to complete. The total numbers of edges and nodes (of any type) generated were 4,512,779 and 4,528,952, respectively. Each distribution network has on average 268 edges and 268 nodes. Despite differing building layouts in suburban test-sites, whole city areas (Newcastle) and the extensive urban zone of London, the synthetic networks created are all visually plausible. Fig 5 and Table 1 compare synthetic with real-world layouts: mismatches show that most actual cables avoid road centre-lines and some avoid roads altogether.

Figure 5: example comparison of (A) detached, (B) semi-detached (duplex), (C) terrace (townhouse), (D) mixed building layouts

Area	Spatial Accuracy (Validated on Feeders)	Mean diff. angle (Validated on Service Lines)
A (Detached)	97 %	17.3°
B (Semi-Detached)	95 %	21.2°
C (Terraces)	96 %	12.6°
D (Mixed Layout)	92 %	28.4°

Table 1: Quantified validation result for sample areas shown in Figure 5

V. Further Work

Further work in refining the algorithm will address

- planarity: ensuring no crossing of edges;
- further information sources: building function is important as establishments such as hospitals and factories have special relationships with the electricity distribution system;
- dependencies: where nodes may be modelled as electricity supply points to water distribution points, such as pumping stations.

The modelling of the latter interactions between different infrastructure networks, and the examination of 'systems of systems' is the ultimate goal of this research.

[1] E. Jenelius, T. Petersen, and L.G. Mattsson, (2006). Importance and exposure in road network vulnerability analysis. *Transp. Res. Part A: Policy and Practice*, 40(7), 537-560.
[2] G. Fu, Z. Kapelan, J. Kasprzyk, and P. Reed (2013). Optimal design of water distribution systems using many-objective visual analytics. *J.Water Resour. Plann. Manage.*, 139(6), 624-633.
[3] R. Chell (2017) Where every drop counts. *Geoconnexion*, July-August, 50-52.