2015-08-29

Establishing a Common Geospatial Approach to Public Transport Service Area Analysis

David O'Connor

Dublin Institute of Technology, dave.oconnor@dit.ie

Follow this and additional works at: http://arrow.dit.ie/beschspart

Part of the Urban Studies and Planning Commons

ESTABLISHING A COMMON GEOSPATIAL APPROACH TO PUBLIC TRANSPORT CATCHMENT AREA MEASUREMENT

Mr. David O'Connor

Lecturer

Dublin Institute of Technology

Abstract

Previous studies (Harrison & O'Connor, ITRN 2012; O'Connor (2), ITRN 2014) have analysed the walking catchment area for bus, light rail and metropolitan rail stops in suburban parts of Dublin city. Public transport users were sampled at each stop and their absolute trip origin identified. This information was then used to identify and approximate the catchment area for public transport at that location.

The purpose of this paper is to collate existing information and establish a common appraisal format using geospatial analysis. Specifically, data from earlier studies will be fed into a geodatabase design and a spatial analytical framework developed for use with ArcMAP 10.2.

This will require geo-referencing of all trip-origins using ArcMAP and coding of each origin-point by mode of travel to/from public transport services. Catchment areas can then be identified using the ArcMAP Network Analyst function. It may be necessary, for example, to distinguish between (i) the core pedestrian catchment and (ii) the extended catchment (affected by range extenders such as cycling, kiss & ride and local feeder services, etc.). Path files will need to be validated against the actual network to eliminate coding, geometric and informal path errors.

A common geospatial approach should yield a more precise measure of the actual extent of public transport catchments across a range of contrasting locations. It should also provide a more robust template for further data capture, a key recommendation of earlier studies.

Database Design for the Dublin Networked Neighbourhood Database

The study objective forms part of a wider research enquiry into *Public Transport Service Levels and the "Networked Neighbourhood"*. The research purpose is to understand the relationship between public transport service levels and neighbourhood catchment areas within urban transport networks. In doing so the study proposes to investigate the potential usefulness of distance travelled to public transport hubs as a revealed preference choice indicator.

In order to collate and analyse data pertaining to service levels and "networked neighbourhoods" (in other words, public transport stops in urban areas), a spatial- or geo-database is required. This document describes the database design required to support the *"Dublin Networked Neighbourhood Database" (DNND)*. The DNND is required to organise and analyse spatial data pertaining to public transport catchment areas and public transport service levels in selected Dublin neighbourhoods.

Design for geo-databases typically comprises the following stages: -
- A systems description
- A list of entities and their attributes
- An entity-relationship diagram (E-R-D)
- Transformation of the ERD into a relational schema
- Layout of survey forms

Systems Description

As part of the wider research surveys of public transport users are taking place at selected public transport stops around the Greater Dublin Area. The purpose of the surveys is to establish the catchment area of public transport stops in contrasting locations and

neighbourhoods around Dublin. It is also intended to understand user perceptions of service levels provided at each location.

In order to establish the catchment area, users are inquired as to their absolute trip origin. The user is prompted for a specific street- or estate-name. Geo-references are required to be generated, based on this information, during data collation and entry.

Users are also inquired as to their mode of arrival to the stop and their mode of onward travel (after using the service). Various other coding data is collected (ticket-type, gender, age-group, etc.). Finally users are asked about their perceptions of service levels on the service they are boarding.

A parallel operational survey, carried out by a second enumerator, accompanies the user survey. At each stop, the number of services during each time period is counted. This yields a headway or frequency. The number of people boarding and alighting at each stop is also counted. There are two critical time periods. Time period 1 is the peak hour. Time period 2 is the shoulder-peak (i.e. the hour immediately post the peak).

DNND Entities and Attributes

The following tables illustrate the geodatabase entities and attributes assigned to the survey data: -

User

User ID	Mode of arrival	Mode of departure	Trip origin (address)	Trip origin (geo-ref)	Gender	Age Group	Ticket-type

User (cont'd)

User ID	Frequency (perceived)	Comfort (perceived)	Convenience (perceived)	Safety (perceived)	Reliability (perceived)

Bus stop

Bus stop number	Number of buses	Passengers boarding	Passengers alighting	Location (address)	Location (georef)	Service type

Time period

Time Period	Time

Enumerator

Initials	Name	Contact email	Contact number

Service type

Service type	Corridor

Entity Relationship Diagram (E-R-D)
The following figure illustrates the geodatabase Entity Relationship Diagram for the Dublin Networked Neighbourhood Database: -
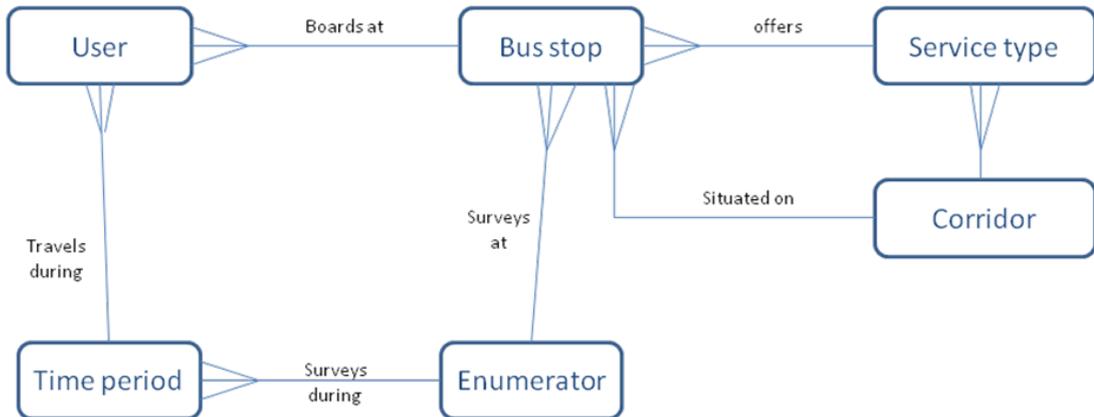
Relational Schema
The following describes the geodatabase relational schema for the Dublin Networked Neighbourhood Database.

Public transport user (<u>User ID</u>, Mode_arrival, Mode_departure, Trip origin (address), Trip origin (geo-ref), Frequency (perceived), Comfort (perceived), Convenience (perceived), Safety (perceived), Reliability (perceived), Gender, Age, Group, Ticket-type*, Bus stop number, Time Period*)

Bus stop (<u>Bus stop number</u>, Number of buses, Passengers boarding, Passengers alighting, Enumerator, Location (address), Location (georef), *Service type, Time Period*)

Time period (<u>Time Period</u>, Time*, bus stop number*)

Enumerator (<u>Initials</u>, Name, Contact email, Contact number*, bus stop number*)

Service type (<u>Service Type</u>, Corridor*, bus stop number*)

Survey Data Collection

Some surveys have already been undertaken – see sample survey sheets in Appendix A – and further surveys are planned). A quota of at minimum 1,000 users is intended. Each stop is chosen to represent a particular corridor- and service-type. A corridor can be, for example, the Malahide Quality Bus Corridor and its environs. A service type can be either arterial QBC (Quality Bus Corridor), HRT (Heavy Rail Transit), LRT (Light Rail Transit), orbital or local. Service types are specific to each stop.

The 2014 studies examined the Stillorgan QBC on Dublin's southside and the Malahide QBC on Dublin's northside within similar urban bands [1, 2]. One orbital stop (on the 17A Dublin Bus route) was included in the Malahide QBC survey. Other sample data was derived from previous studies. 2012 studies looked at rail stops in four bands across the urban area, including: Urban, Outer Urban, Inner Suburban and Outer Suburban [3]. The number of stops and surveys collected is shown in the table below.

	Date	Service Type	Number of Stops	Number of Users
Malahide QBC	April 2014	QBC	4 stops	177
17A	April 2014	Orbital service	1 stop	17
Stillorgan QBC	June 2014	QBC	4 stops	139
DART	March 2012	HRT	4 stops	200
LUAS	March 2012	LRT	4 stops	200
Total			**17 stops**	**733**

A total of 733 public transport users have thus far been surveyed across 17 public transport stops. Some inconsistencies in data collection methodology exist, which is a potential limitation of the research. For example the DART and LUAS surveys collected data from walkers only.

Coding of Locations

This paper uses the Stillorgan bus stop (bus stop #7474) on the Stillorgan QBC data as a test case for coding of locations. The figure below shows the catchment area for all users surveyed at the bus stop by mode of arrival, as coded into ArcMAP 10.2. Each datapoint contains a range of attribute data, such as mode_of_arrival, gender, Comfort_perceived, etc. Each datapoint can also be related to entities, such as Bus_stop, Time_period, etc.
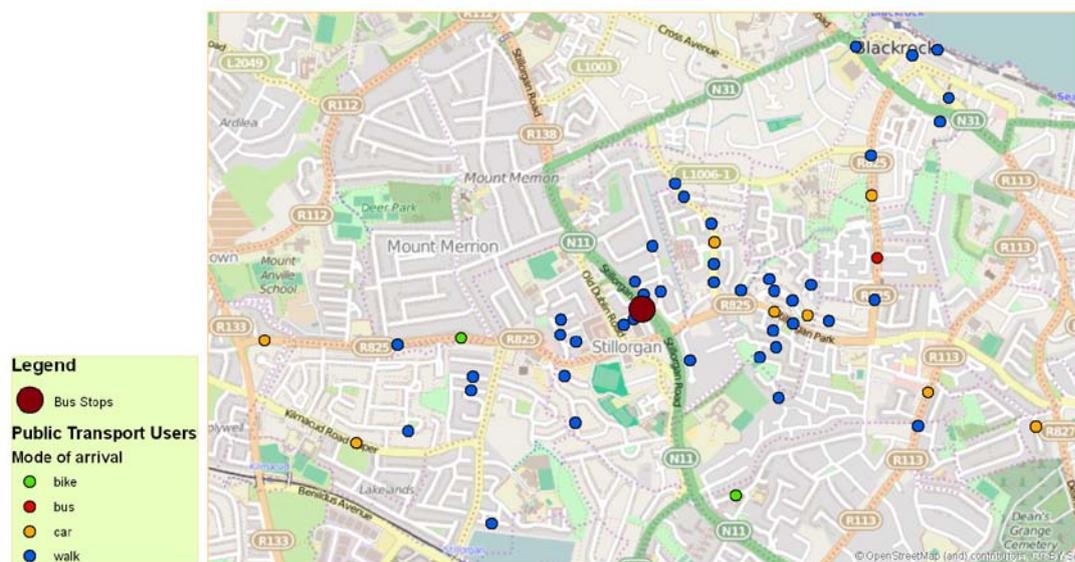
Further Research Stages and Uses of the DNND

Once all data points are coded and trip origins geo-referenced, a *network dataset* is required. A network dataset is a dataset that describes the travelling of an agent - e.g., pedestrian, truck driver, cyclist, etc. - along a route [4]. This will allow distances to be measured between trip origins and the respective public transport stops. Aggregated distances can be calculated by mode. This should help to identify, with greater precision, both the potential catchment 'shed' and, more elementally, in establishing a cordon area principle or protocol. This may be based, for example, on use of the 85th percentile or other appropriate means.

The DNND facilitates the measurement of the propensity to travel to public transport. Once this is understood it should be possible to relate this to other geo-spatial attributes within the respective catchment areas, a key part of the wider research framework. Relevant attributes can be derived from a variety of data sources, some of which identified in the following table.

Census Small Area Population Statistics	*Socio-economic status*
	Age / occupation / education levels
	Self assessed health levels
	Mode of travel to work
	Distance travelled to work
	Age of house
Open Street Map	*Permeability index*
	Density
	Area
	Open space
	Gradient
	Adjacent public transport corridors
Dublinked Store	*Adjacent public transport corridors*
	Level of service
POWSCAR	*Trip distributions*
DNND Operational Survey	*Level of service*

Data and Study Limitations

Open Street Map (OSM) data is comprehensively available in the areas of interest. Limits are applied within OSM to the size of network datasets which can be created [5]. A significant stage will be to establish whether OSM can be utilised to analyse the catchment areas likely to be generated by the research data or if an alternate (non-open source) data source is required.

The Dublinked datastore contains up-to-date public transport level of service and infrastructure data. A critical open data platform, Dublinked publishes most of its data in MapInfo-ready format. MapInfo is a commercial GIS corporate product produced by Pitney Bowes. The application used for this study was ArcInfo, a comparable and industry-standard product produced by the Esri Corporation. ArcInfo will read MapInfo files natively if the Data Interoperability Extension is installed. Alternately, applications such as ogr2gui, FME Tools and opensource platforms such as QGIS, provide file conversion facilities which may be effective.

Some inconsistencies in data collection methodology exist across the various surveys. For example the DART and LUAS surveys collected data from walkers only. Ticket-type and age data is omitted from some surveys.

Also the level of precision in coding of trip-origins varies according to the information from respondents. Interviewees are asked for their street or residential estate of origin. Frequently, travellers provided the names of estates or districts, rather than street names. In either case the trip was geo-referenced approximately around the street or estate centroid and plotted accordingly.

A key decision pertains to the inconsistencies and precision levels identified above. It will need to be decided whether the margins of error are acceptable or if a "clean" dataset needs to be generated. This will be easier to understand once the full dataset has been coded and presented for initial observation.

Conclusions

This paper describes the design and system objectives for the "Dublin Networked Neighbourhood Database" (DNND). The purpose of the DNND is to support research into public transport catchment areas and the extent to which these are influenced by levels of service and other factors. Initial survey data has been collected and processed satisfactorily through the DNND. Subsequent research will entail creating a network dataset and cross-analysing existing public transport user behaviour with relevant attributes and datasets from external sources. An extensive review of literature is also entailed.

References
[1] O'Connor, D. Stillorgan QBC Public Transport Level of Service Analysis, Proceedings of the Irish Transport Research Network 2014
[2] O'Connor, D. Malahide QBC Public Transport Level of Service Analysis, Proceedings of the Irish Transport Research Network 2014
 [3] Harrison, O. & O'Connor, D. Measuring Rail Catchment Areas in the Greater Dublin Area, Proceedings of the Irish Transport Research Network 2012
[4] ArcGIS Resources, What is a Network Dataset?
http://resources.arcgis.com/en/help/main/10.1/index.html#/What_is_a_network_dataset/0047 00000007000000 accessed on 20th August 2015
[5] Github, Create a Network Dataset from OSM Data, https://github.com/Esri/arcgis-osm-editor/wiki/Create-a-network-dataset-from-osm-data accessed on 20th August 2015

Appendix A: sample of survey questionnaire

Dublin Institute of Technology / Northside Partnership
Public Transport Catchment Survey
Malahide QBC & environs, 10th April 2014
Student Transport & Mobility Assignment

Hi, would you have one or two minutes to help with a survey? We are students of spatial planning at DIT and we are doing a study with Northside Partnership (a local community development organisation) to see how well public transport (excluding the Schools Service) is used in the area.

PART A — How did you get to this bus stop today?

Walk ☐ Bike ☐ Car ☐ Bus ☐ Other (state) _____

Where did you start your journey (can you give us the street / estate name or an appoximate location)?

Location: _____

How will you complete your onward journey today (please state the principal mode)?

Walk ☐ Bike ☐ Dublinbike ☐ Bus ☐ LUAS ☐ DART ☐ Rail (other) ☐ Car ☐ Other (state) _____

PART B — Would you also have time to tell us (where 1 is very bad and 10 is very good) how you would rate the service you are taking in terms of...

	1	2	3	4	5	6	7	8	9	10
Frequency?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Comfort?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Convenience?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Safety?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Reliability?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Enumerator use only
Gender of interviewee (circle one): Male Female
Enumerator (initials):
Location (circle one): Clare Hall Coolock Artane Priorswood OTR
Survey period (circle one): 08:00 - 09:00 09:00 - 10:00

	Period 1: 08:00 - 09:00	Period 2: 09:00 - 10:00
Buses		
Passengers boarding		
Passengers alighting		

Enumerator use
Enumerator (initials):
Location: