Timetable segment mapping process

Scott Mackenzie

September 1, 2014

1 Steps

1. Run the macro "Tiploc List 1" on the sheet "Route definitions" over the segment data to be mapped. This will produce a sheet containing a list of all TIPLOCs in the data. Look at TIPLOCs that are only present in one segment file and add them to the reverse directed file if required.

2. Run the macro "Segment mappings R" on the sheet "Segment Mappings". This will produce a file *SegmentMappings_R.txt*.

3. In the R script *ProcessTimetables_v3.R* set as the main directory (MainDir) the location where you want all output files to be generated. Set the timetable directory (TimetableDir) to be the location where the XML timetable files are located. Place the file *SegmentMappings_R.txt* in the main directory.

4. Run the script *ProcessTimetables_v3.R*. This will produce two tab delimited output files, *Timetable_TIPLOCs.txt* and *Timetables.txt*. The file *Timetables.txt* contains the following fields:

   - **Timetable number**
   - **Headcode**
   - **Days of operation**. The days of the week which this timetable is run.
   - **Start date**. Date from which the timetable is in operation.
   - **End date**. Last date which the timetable is valid.
   - **Timetable start time**.
   - **Timetable end time**.
   - **Start TIPLOC code**.
   - **Finish TIPLOC code**.

   The file *Timetable_TIPLOCs.txt* contains the following fields:

   - **Timetable number**.

- **Headcode**.
- **TIPLOC description**.
- **TIPLOC Code**.
- **TIPLOC time**.
- **Line code**.
- **Platform**.
- **TIPLOC number**. This will initially be set equal to "NA".
- **TIPLOC mapped**. This will initially be set equal to "–".

5. In the script *GenerateData_v3.R* once again set the main directory as above and set the segment directory (SegDir) to where all the segment files (*.seg*) are located. This script will produce a folder "Mapping Data" in the main directory. In the file *Timetable_TIPLOCs.txt* TIPLOCs will now be assigned a number in the "TIPLOC number" field. TIPLOCs that remain "NA" are ones that are not present in the route data.

6. Set the main directory in the script *ShortestPaths_v3.R* and run it. This will produce the files *ShortestPreviousNode.txt* and *DistMatrix.txt* in the folder "Mapping Data". This script calculates the shortest path between every pair of connected segments in the network.

7. Set the main directory in the script *BuildRoutes_v3.txt* and run it. The last column in the file *Timetable_TIPLOCs.txt* will be set equal to "TRUE" if the TIPLOC has been mapped. A file *TTRoutes.txt* will be generated listing the segments that have been selected to map the timetable. The TIPLOCs that have been mapped are shown next to the corresponding segment file they were found on. The fields in this file are:

- **Timetable number**.
- **Headcode**.
- **segment file**
- **Number TIPLOCs**. This is the number of TIPLOCs from the timetable found on the segment.
- **TIPLOC descriptions**. The remaining columns are the TIPLOCs that were found on the segment in the order that they were found.

2 Dependencies

The script *ProcessTimetables_v3.R* need only be run once at the beginning of the segment mapping process to strip the XML timetables of unwanted characters and fields.

ShortestPaths_v3.R only needs to be executed when changes are made to the segment mapping file *SegmentMappings_R.txt*. If the only changes to the data

are the addition of TIPLOCs then it is not necessary to recalculate the shortest paths.

The script *GenerateData_v3.R* must be run after either TIPLOC or segment mapping data changes are made.

3 Interpreting output

If the script produces a path connecting TIPLOCs $A \rightarrow B \rightarrow C$ then it is guaranteed to be the shortest path that connects these three TIPLOCs. It is possible that there is a shorter path between A and B, or B and C but not A to C passing through B.

If there is no path connecting adjacent TIPLOCs then this will be indicated by an empty field in the segment file column in *TTRoutes.txt*. This could be the result of one or more of the reasons below:

- Missing TIPLOCs

- Missing track data

- Missing or incorrect mappings

If there is a large number of segment files connecting adjacent TIPLOCs then this is likely to be incorrect. This could be the result of any of the reasons above.