CyIPT Manual

Version 0.1 Friday, 09 March 2018

Contents

Introduction

The Cycling Infrastructure Prioritisation Toolkit (CyIPT) is a research project based at the University of Leeds and funded by the Department for Transport (DFT). The purpose of CyIPT is to develop methods and tools to assist in the design and planning of new cycling infrastructure. CyIPT is currently (as of March 2018) a working prototype. Therefore, any recommendations produced by CyIPT should be subjected to independent assessment before making investment decisions.

CyIPT Team

Dr Robin Lovelace, University of Leeds

Dr Malcolm Morgan, University of Leeds

Prof John Parking, University of the West of England

Martin Lucas-Smith, CycleStreets

How CyIPT works

This section gives an overview of how CyIPT works and some of its main limitations. For full details, see the Technical Details section below.

Figure 1 outlines the basic structure of CyIPT. First CyIPT takes data about each road and path in England and uses it to recommend the approximate type of cycling infrastructure. The recommendation algorithm is based on Highways England's Interim Advice Note 195/16. The algorithm can make eight possible recommendations (Cycle Lanes, Cycle Lanes with light segregation, Cycle Street, Cycle Land on Path, Stepped Cycle Tracks, Segregated Cycle Track on Path, Segregated Cycle Track, and None)
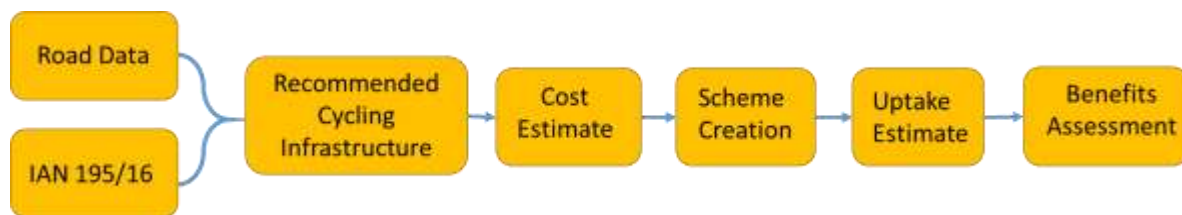
Figure 1: Basic CyIPT Workflow

Based on the length and type of infrastructure, CyIPT estimate the cost of constructing the recommended new cycling infrastructure. CyIPT take account of existing infrastructure and therefore does not apply a cost of building existing cycle infrastructure if it is of sufficient quality.

As CyIPT's recommendations are made for each road segment (junction to junction), they can be for very short sections of road. Therefore, CyIPT has a clustering algorithm, which attempts to take recommendations and group them into coherent schemes that could be constructed. One the schemes have been produced, CyITP estimate the number of additional cyclists the scheme would produce and performed a benefit cost assessment of the schemes to identify the schemes which are most likely to be worth building.

The CyIPT process is deterministic, which is to say that it produces the same results each time the model is run, and has therefore been pre-processed by the CyIPT team. The results for England are made available through the CyIPT website (www.cyipt.bike). The CyIPT website also allows for data download and the source code is available at GitHub (https://github.com/cyipt).

Key Concepts

Rate Limiting

To maintain the performance of the CyIPT website, the map data is rate limited. These limits will prevent you from downloading too much data at once. You are most likely to encounter rate limiting when attempting to view large areas. When rate limiting occurs, the data will appear on the map in a patchy or random fashion.
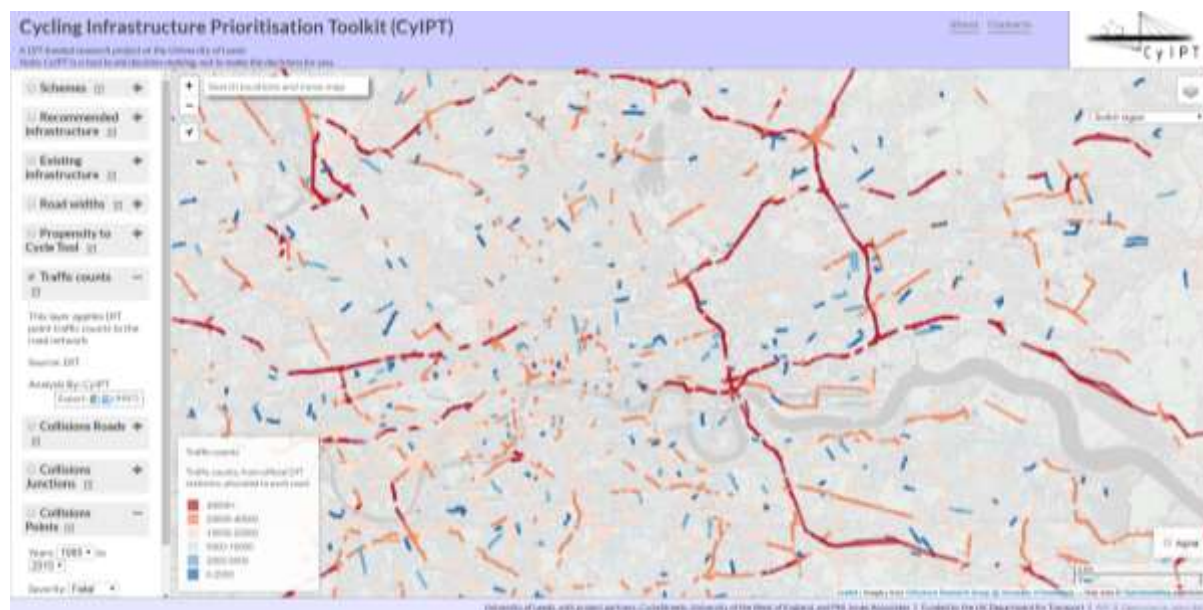

Figure 2: Example of the map of London when rate limited is in effect

For some layers now data will show until you are zoomed in to a certain level. While some layers (e.g. Propensity to Cycle Tool layer) provide more detail the further you zoom in.

Layer Descriptions

The CyIPT website is an interactive map presenting a range of different

Schemes

Recommended infrastructure

The recommended infrastructure layer presents CyIPT's recommendation for the type of cycle infrastructure required on each road segment. Roads and paths where no cycle infrastructure is required are not show. The recommendations do not take account of any existing infrastructure; therefore, the recommendations may suggest that existing cycle infrastructure remain, be upgraded, or in limited cases be downgraded.

It is important to understand that the CyIPT recommendations are based on cycling levels in the 2011 census, and it may therefore underrepresent current cycling levels and infrastructure requirements. For more detail on the recommendation process, see the Technical Details section below.

Existing infrastructure

The Existing Infrastructure layer has three modes (Cycle infrastructure, Speed limits, Footways). These modes are intended to help the users understand how CyIPT 'sees' the world. For each mode, the underlying data is taken from the Open Street Map (OSM), which is then passed through a cleaning process to correct errors and fill in any missing data.

If any values are incorrect in CyIPT (for example an incorrect speed limit) clicking on the road will show a popup including an "Edit in OSM" button. This button will take you to the same road in the OSM website where you can make corrections. These corrections (and any other updates) will be incorporated into CyIPT next time the data is rebuilt.

Cycle infrastructure

The cycle infrastructure mode shows physical cycling infrastructure such as cycle lanes and tracks. It does not show signposted cycle routes (e.g. national cycle network) where there is no physical infrastructure.

Speed limits

The speed limits mode shows the maximum speed limit in miles per hour.

Footways

The footways mode show the presence of footways (pavements) on the road.

Road widths

The road width layer has two modes (road width, width status).

Road width

The road width layer shows and estimate of the road with based on the OSM. The calculation use standard lane widths and information on the number of lanes and footways to estimate the total width of the carriageway.

Width status

The width status layer compares the estimated width of the road with any proposed new cycling infrastructure to width of the road as measured from the Ordnance Survey MaterMap. This mode helps highlight locations where road space is limited and space reallocation may be required to provide new cycling infrastructure.

Propensity to Cycle Tool

The Propensity to Cycle Tool layer takes date from the Propensity to Cycle Tool (PCT) (www.pct.bike) and reanalyses for CyIPT. The layer is analogous to the PCT LSOA route network layer, which can also be accessed as a base map though CyIPT. The difference with the PCT version is that the CyIPT version is matched to the OSM roads and is therefore both more detailed and interactive.

The PCT layer supports the same 5 scenarios as the main PCT tool, however only the 2011 census scenario is used within CyIPT.

Traffic counts

The traffic counts layer shows CyIPT understanding of road traffic levels. This comes from the DFT traffic count data, which CyIPT then matches to the road network. As the DFT data is not provided for all locations in all years, CyIPT takes the most recent available year. The DFT data has good coverage of the strategic road network, but more limited coverage of minor roads.

Collisions Roads

This layer should be used in conjunction with the Collison Junctions layer.

This layer takes the Stats19 data on road collision between 1985 and 2015 and matched collisions that did not occur at a junction to the nearest road. This then provides a count of the total number of collisions that occurred on that road. These can be filtered by the severity of collisions (slight, serious, fatal) and severity of bike casualties (slight, serious, fatal).

Collisions Junctions

This layer should be used in conjunction with the Collison Roads layer.

This layer takes the Stats19 data on road collision between 1985 and 2015 and matched collisions that did occurred at a junction to the nearest junction. This then provides a count of the total number of collisions that occurred on that road. These can be filtered by the severity of collisions (slight, serious, fatal) and severity of bike casualties (slight, serious, fatal).

Collisions Points

This lay shows the locations of all collisions in the Stats19 data from 1985 to 2015, it can be filtered by year and severity.

Technical Details

CyITP is developed in the R Programming language and is an open source project. The code is available at https://github.com/cyipt. CyIPT consists of two main parts; the R based analysis code, and the website, which is a mixture of HTML/CSS/JavaScript with a PostgreSQL database accessed via a PHP-based API.

This manual focusses on the R based analysis code, as the website exists purely for visualisation and ease of use.

Data Preparation

CyIPT is reliant on some pre-existing datasets from third parties. While many of these are publically available, CyIPT required them to be pre-processed before use. These scripts are provided for context, but in most cases, users should download the pre-processed data directly from GitHub.

CyIPT Master Script

The CyIPT master script https://github.com/cyipt/cyipt/blob/master/scripts/cyipt.R can be used to run the whole CyIPT process. It manages several global settings

Settings

skip

Should the code skip regions that have already been done?

overwrite

Some stages overwrite existing files, for example by adding an extra column of data

Note that not overwriting may cause later stages to fail if they expect earlier stages results to be in the starting file

ncores

Some functions use parallel processing how many clusters should be run? This should be less than the number of cores on your computer.

verbose

Get extra messages and information while CyIPT is running.

all.regions

Ignore the regions to do file and run for all regions.

Regions to Do

If all.regions = FALSE CyIPT will choose which regions to run for based on the RegionsToDo file
https://github.com/cyipt/cyipt/blob/master/input-data/RegionsToDo.csv

Simply place a y in the do column of this csv file next to regions that you wish to run.

CyIPT uses the 2011 travel to work areas produces by the Office for National Statistics (ONS). However it could be changes to accept a different se to regions.

Libraries

CyIPT requires the following R libraries to be installed prior to use. CyIPT was developed for R 3.4.0.

sf
osmdata
stringr
dplyr
parallel
xgboost
igraph
tmap

Step 1: Download the Data

https://github.com/cyipt/cyipt/blob/master/scripts/prep_data/download-osm.R

This script downloads the OSM road and path network for each region.

Inputs

Regions.todo

Boundaries file "../cyipt-bigdata/boundaries/TTWA/TTWA_England.Rds"

Outputs

Region Boundaries "../cyipt-bigdata/osm-raw/",region_nm,"/bounds.Rds"

OSM road network "../cyipt-bigdata/osm-raw/",region_nm,"/osm-lines.Rds"

OSM road junction points "../cyipt-bigdata/osm-raw/",region_nm,"/osm-junction-points.Rds"

Parallelised

Yes

Step 2: Clean the OSM Tags

https://github.com/cyipt/cyipt/blob/master/scripts/prep_data/clean_osm.R

This script "cleans" the OSM data by removing or correcting errors and filling in missing data with best guesses. Guessing is required, as some later stages of CyIPT require information (such as speed limits) which is not always available. In isolated cases of incorrect guesses it is best to correct the value in the OSM using the "edit in the OSM" button on the CyIPT website. These corrections will then be incorporated into the next build of CyIPT. In general cases of CyIPT miss-guessing, please submit and issue via GitHub https://github.com/cyipt/cyipt/issues

Inputs

Regions.todo

OSM road network "../cyipt-bigdata/osm-raw/",region_nm,"/osm-lines.Rds"

Outputs

OSM road network "../cyipt-bigdata/osm-clean/",region_nm,"/osm-lines.Rds"

Parallelised

No

Detail

The cleaning process consist of the following stages

1. Removing un-allowed road types (e.g. planned or demolished)
2. Replacing depreciated highway tags
3. Cleaning the junction tag
4. Summarising the one-way nature of roads
5. Guessing the max speed of roads with an unknown max speed based on road type
6. Guessing the presence of footways (pavements) with unknown footway status
7. Summarising the presence of bridges and tunnels
8. Cleaning the segregation status of cycle infrastructure
9. Cleating a summary of the road type
10. Cleaning and/or guessing the number and nature of lanes in each direction
11. Cleaning the tagging of cycle infrastructure and improving detail of what is on each side of the road.

Step 3: Get traffic counts

https://github.com/cyipt/cyipt/blob/master/scripts/prep_data/get_traffic.R

This script assigns the point traffic count data to the road network.

Inputs

Regions.todo

OSM road network "../cyipt-bigdata/osm-clean/",region_nm,"/osm-lines.Rds"

Traffic Points "../cyipt-bigdata/traffic/traffic.Rds"

Outputs

OSM road network "../cyipt-bigdata/osm-clean/",region_nm,"/osm-lines.Rds"

Parallelised

No

Details

This scrip divides the point traffic counts based on whether they are on classified (e.g. M21, B340) or unclassified roads. Unclassified road points are matched to the nearest road in the OSM, and therefore the value only extends a short distance from the point location. For classified roads, a series of Voronoi polygons are constructed around the points and all the road segments within each polygon are assigned the value of their nearest point. This provides continuous coverage, but can produce some erroneous results such as off ramps having the same traffic levels as the main carriageways.

In both cases, the script takes the Annual Average Daily Traffic (AADT) value from the most recent available year. For the strategic road network, data is mostly from 2015/2016, but for minor roads, it can be significantly earlier. For the purposes of CyIPT is the traffic data is mostly used for identify the very busy and most hostile roads, thus this inconstancy of data is not a significant problem. However, users intending to use the data or method for other purposes should consider the implications of this inconstancy within the data.

Step 4: Split the lines at each junction

https://github.com/cyipt/cyipt/blob/master/scripts/prep_data/prep_osm.R

This script splits the roads at each junction into road segments.

Inputs

Regions.todo

OSM road network "../cyipt-bigdata/osm-clean/",region_nm,"/osm-lines.Rds"

OSM road junction points "../cyipt-bigdata/osm-raw/",region_nm,"/osm-junction-points.Rds"

Outputs

OSM road network "../cyipt-bigdata/osm-prep/",region_nm,"/osm-lines.Rds"

Parallelised

No

Details

The splitting of the roads at junctions is mostly required for the later application of the PCT data. Within the OSM a road may be represented by a single long line crossing several junctions. However, at each junction cyclists may join or leave the road. Therefore, it is not appropriate to analyse the road network as it is represented in the OSM. By splitting the road network into segments it ensures that, the analysis is appropriately detailed.

Note the splitting is done by cutting tiny holes out of the road lines (r = 0.01m) therefore the lines are no longer touching; this would prevent this dataset being used in a routing engine.

Step 5: Get the PCT estimate of number of cyclists
source("scripts/prep_data/get_pct.R")

Step 6: Get road width estimates
source("scripts/prep_data/get_widths.R")

Step 7: Get Collisions
source("scripts/prep_data/get_collisions.R")

Step 7: Evaluate Infrastructure Options
source("scripts/select_infra/select_infra.R")

Step 8: Compare Widths Needed to Widths Available
source("scripts/select_infra/compare_widths.R")

Step 9: Group into schemes
source("scripts/select_infra/make_schemes2.R")

Step 10: get uptake
source("scripts/uptake/calc_uptake_routechange3.R")

Step 11: Calculate Benefits
source("scripts/benefits/eval_benefits4.R")

General Comments
This section covers general comments and warnings about CyIPT.

UK Bias
CyIPT was developed for England and has several UK biases built in. For example:

1. All geospatial analysis is perfumed using the British National Grid
2. Left had driving is assumed

3. Speed limits are in mph

If you wish to use CyIPT outside the UK these assumption will not hold and need to be fixed in the code.