

Position Paper: Open Data-Logging for Bikes

Linking Geospatial Data W3C workshop - <http://w3.org/2014/03/lgd/>

Rich Boakes, Jacek Kopecky & Oliver Williams

rjb@port.ac.uk, jacek.kopecky@port.ac.uk, cam00896@myport.ac.uk

January 19, 2014

Abstract: Cycling in UK cities is often perceived as unsafe, yet there is no useful empirical data recording cyclists experiences. We are working on a project to develop specialised sensor instrumentation for bicycles that will record traffic data automatically, and publish aggregate information, for instance to support analysis by local governments or cycling clubs.

Problem

Cycling in UK cities is not perceived as safe by many people. Those who *do* cycle on the roads are often perceived as taking a big risk. Many people¹ who would *like* to try cycling are *either* put-off by horror stories in the press, *or* if they do venture onto the roads, vehicles passing closer than is comfortable can quickly give the impression that cycling is frightening and dangerous. Journeys that could easily be made by bike are thus made by car, compounding the problem.

Many cyclists have taken to wearing helmet cameras, however video evidence of near misses (e.g. <http://youtu.be/v2NBcIFD15s>) may add to the public perception of danger rather than act as a catalyst for change.

When local councils spend money on making roads safer for cycling, their investments are not necessarily well targeted and there is often no way of gathering evidence to assess if spending has achieved its aims. Such evidence *should* make future decisions and investments easier to justify and better targeted.

Hypotheses

To understand whether a novel form of data collection can:

1. identify places where cyclists are most at risk from vehicles in close proximity.
2. identify places where cyclists feel most at risk.
3. identify any disparity between actual and perceived risk in cyclists.

If successful it is anticipated that such data collection methods could be used by local councils to target areas for improvement, and provide a baseline for measurement of success. It is hoped this may catalyse a shift to more sustainable forms of transport.

¹ Derived from an undergraduate research questionnaire with 2.5K respondents, to be published end-of-term 2014.

Proposed Solution

We recognised that gathering empirical data would require the instrumentation of a bike and proposed to use off-the-shelf hardware to create a prototype telemetry system.

Prototype One

A first prototype was created in March 2013. We used a lot of rubber bands, foam, a Raspberry Pi, a GPS unit, a battery pack, and a single ultrasonic sensor attached to the seatpost.



fig1. The First Prototype

During these tests we also used helmet cameras to review rides and understand the data we were seeing. The data we collected from individual journeys was processed and visualised as a heatmap - this is the first map we successfully created in April 2013.

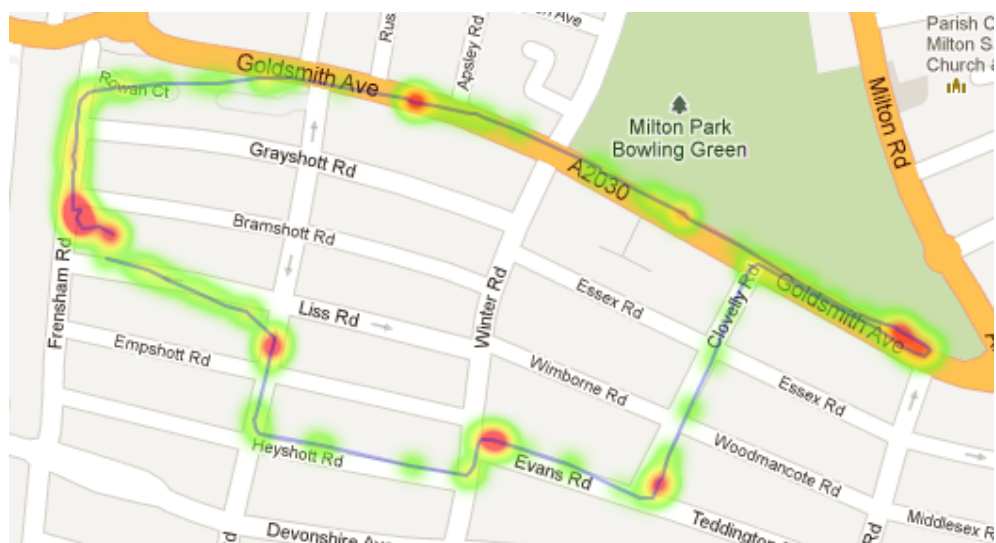


fig2. The First Heatmap

Prototype Two

The second prototype improves on the first in several areas.

1. We've started using the GeoJSON format for logs that are uploaded to the server.
2. We're beginning to turn it into a unit that can be more easily used to log data (and not just prove that logging is possible). To this end a physical interface (a button) has been added that enables logging to be started and stopped.
3. We are extending the unit to use more sensors:
 - a. left right and rear-facing ultrasonics,
 - b. rear facing camera.
4. We are building web services for capturing data from multiple logging units.
5. We are investigating data users may wish to query and use the data, so that we can structure it appropriately, and build suitable APIs for its access.

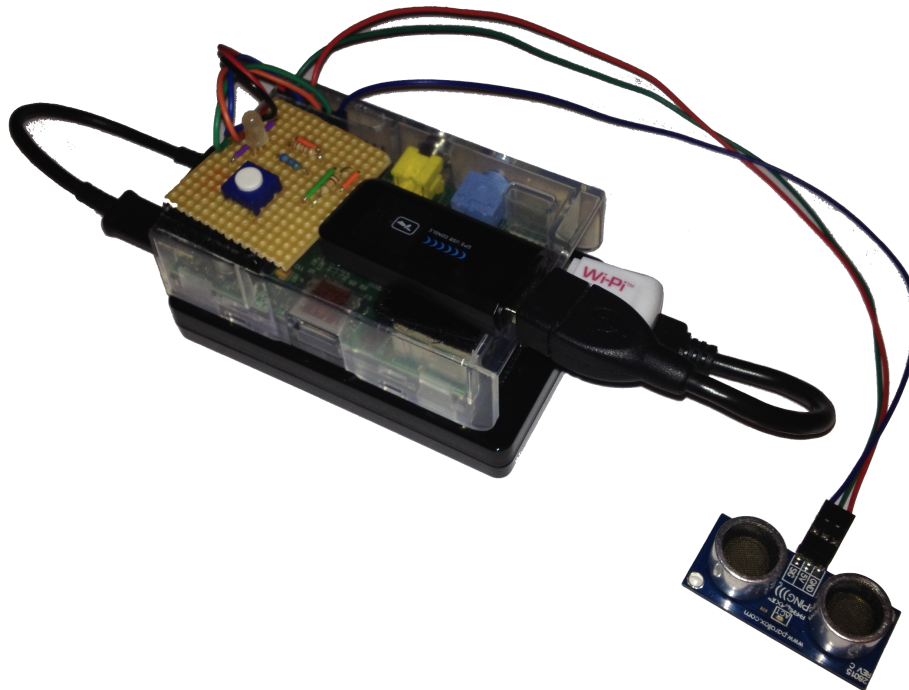


fig3. The Second Prototype

Our aim is to run trials along specific roads in Portsmouth over the next 6 months, in conjunction with Portsmouth Cycle Forum (the local cycle enthusiast & lobbying group). We aim to collect and collate data from thousands of cycle journeys and build an aggregate proximity heatmap. The volume of journeys should overcome one-off blips and show the areas where close proximity occurs most regularly.

We think a map such as this might be one of the useful outputs of such a project. It should help local councils to target road improvements. During and after any subsequent 'improvements' to the road, telemetry measurements will continue to be taken, so that the effects of any changes can be visualised and discussed. This is one example, and we expect there may be more.

Next Steps

Our initial prototype work has convinced us that the collection, collation and free availability of cycle telemetry data could provide valuable evidence in understanding how the design of roads (a) may place vulnerable road users in danger, or (b) may give them the perception of danger, and thus limit or reduce the use of sustainable transport. We also expect that the data we are proposing to gather may provide insight (that is currently available) into the behaviour of cyclists and other vehicle users.

The next version of the prototype is may contain:

- An improved physical interface enabling cyclists to record moments when they are frightened (to actively record perception and not just factual data)..
- Barometer and motion sensors.
On-board processing of image data to include OCR for number-plate recognition.
- Live audible and/or visual feedback of data to the cyclist, e.g. a read-out of just how close the last near-miss was.

Workshop Interest

We intend to continue developing the hardware and software² and welcome ideas and contributions from others. We have experience with a linked geo-data project ParkJam³ which imported data about car park locations from OpenStreetMap via the LinkedGeoData project, and now publishes live crowdsourced parking availability data (admittedly, with little crowd usage at this time) and we are minded to adopt similar technologies going forward. The opportunity to discuss the project with such an appropriate and informed community is extremely desirable.

In the workshop, we would be looking to discuss mainly the following points:

- publishing traffic heatmap and incident data in a geo-linked form,
- handling privacy concerns for the data published (including whether we should use OCR to identify and remove number-plates, or to identify dangerous drivers; and also whether individual cyclist submitters could somehow be identified, and how to avoid this),
- managing an open API for data access and for submissions from project participants (instrumented bicyclists).

² Software is hosted on GitHub at <https://github.com/Oliverr/PicycleApp> and <http://oliverr.github.io/PicycleApp/>

³ <http://parking.kmi.open.ac.uk/>