

e-Science in the Streets: Urban Pollution Monitoring

Anthony Steed¹, Salvatore Spinello¹, Ben Croxford², Chris Greenhalgh³

¹Department of Computer Science, University College London

²Bartlett School of Architecture, University College London

³School of Computer Science & IT, University of Nottingham

ABSTRACT

Current pollution monitoring systems use small numbers of fixed sensors that report multiple species of pollutant. In the UK, online archives of such monitoring sites go back over 30 years. Such archives give a detailed picture of trends in pollution and characteristics of specific pollutants such as average levels at different time-scales. However pollution levels can vary at a small scale. For example, carbon monoxide is concentrated around transport routes and disperses rapidly over a few tens of metres.

This project investigates local variations in pollution by using mobile sensors. These sensors aren't as well calibrated as fixed sensors, but they do give a spatial picture of pollution variation. We are integrating readings from multiple mobile sensors to create maps of pollution variation at a per-street level in order to enable new kinds of investigation of pollution. We are also planning to investigate remote collaboration between scientists in the field and in the lab by using a collaborative visualisation system.

1 INTRODUCTION

In their everyday life, most citizens in the UK are rarely concerned with the effects of urban pollution. However, high concentrations or extended periods of exposure can have serious health effects. Many local councils monitor urban pollution to determine, for example, the effects of local traffic congestion. However this is usually done from a small number of sparsely distributed, though highly accurate, pollution sensors. For the average person travelling through the city, such numbers are difficult to interpret because a single value might represent a sample several square kilometres away. However it is known that pollution levels can vary, sometimes dramatically, on a per street basis.

In the EQUATOR IRC e-Science project Advanced Grid Interfaces for Environmental Science in the Lab and in the Field (EPSRC grant GR/R81985/01) [1], we have been investigating an alternate way of mapping pollution using tracked mobile sensors. An accurate carbon monoxide sensor is coupled with a GPS receiver and a logging device. This device can be carried or placed on a

bike rack. The device logs time, position and pollution level. The resulting recordings are less accurate, but from a wider area of samples. With several such devices being carried around, it will be possible to build a map that shows detailed local variations in pollution.

1.1 Air Quality

The pollutant we are studying is carbon monoxide. Transport makes the greatest contribution to carbon monoxide levels and carbon monoxide affects urban areas more significantly than rural areas. Overall carbon monoxide levels have fallen since the 1970s, averaging $1\text{mg}/\text{m}^3$ [2].

The Air Quality Site contains archive data from over 1500 UK monitoring stations going back in some instances to 1972 [3]. Such data sources give a good picture of variation from urban to rural areas. In urban areas some sense of potential variation is conveyed by the difference in readings between kerbside sensors and sensors placed in background areas away from pollutant sources. However they don't capture the detail of per street variation.

Carbon monoxide disperses over a matter of hours, but Croxford et al. have shown that this is affected by local street configuration [4]. This study used a cluster of sensors in fixed placements in a small area around University College London (UCL). The Air Quality Strategy for England, Scotland, Wales and Northern Ireland [5], suggests a standard of 10ppm ($11.6\text{mg}/\text{m}^3$) running 8-hour mean. In the vicinity of UCL, the Croxford study found a peak CO concentration of 12ppm, but nearby sensors reported much lower values near the background level for CO. Thus, moving pedestrians or vehicles would probably not experience this peak for a long period.

Our approach to bridging this gap between large, sparse and small, dense studies is to use tracked mobile sensors that are carried by pedestrians or mounted on vehicles. We have built a prototype and are completing the infrastructure to dynamically build maps of pollution around UCL. There are several related works in the UK that aim to increase understanding of pollution variation. These include dense sensors network of the Discovery



Figure 1 Early prototype of the mobile pollution-logging device

Net e-Science project [6], and the combined sensing and air flow modelling approach of DAPPLE [7].

2 SYSTEM OVERVIEW

At the current stage we have built five devices based around a design that is simple to construct. We are investigating two infrastructures for interacting with the data. The first is a data-logging mode where the data is captured to a database. The second is a collaborative mode, where pollution is integrated as another resource for users to share and discuss whilst visiting an area

2.1 Components

To make the design flexible we based the first prototype

around a HP Jornada 568 handheld computer, see Figure 1. This is connected to a GPS receiver and a pollution sensor. The GPS receiver was originally a Garmin Etrex Vista, a consumer handheld unit as shown in Figure 1. This has since replaced with a dedicated, screen-less part also manufactured by Garmin. The pollution sensor is manufactured by Learian Design Ltd [8].

The device can record position as WGS84 latitude, longitude and height, and pollution and temperature every second. The GPS receiver also supplies accurate time for the logs. The device has a real-time display of pollution data and satellite information. The device has no built in wireless networking. Thus with the current prototype, data upload is only possible when the device is synchronised with a desktop.

Two second-generation prototypes are being constructed at the time of writing. The first is a dedicated data logger that integrates all the sensors into a custom package with a single battery. It will have roughly the same user functionality as the first prototype since it has a two-line LCD screen and it will support explicit data upload only. This is designed to better support the data-logging mode described in Section 2.2.

The second second-generation prototype is based around the Bitsy/Wearable platform that is being used in the companion e-Science projects in Equator [9][10]. It will support wireless networking so that data upload can be done in a proactive manner whenever the device can detect the relevant infrastructure.

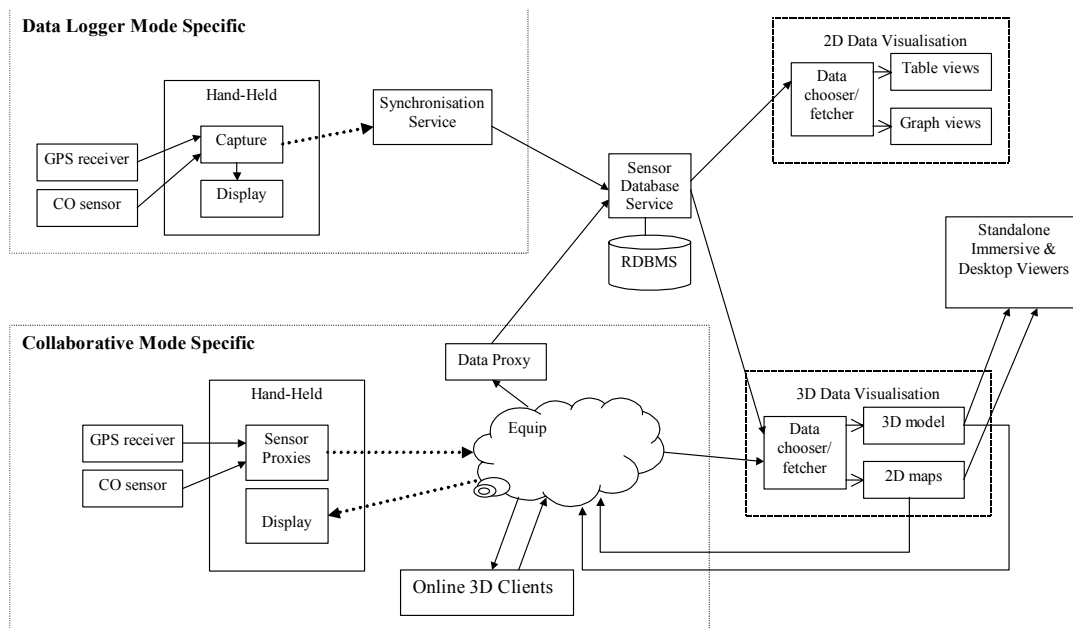


Figure 2 Overview of system architectures for data logging mode and collaborative mode



Figure 1 Original NTF data

2.2 Data Logging Mode

To support visualisation applications the most important functions of the devices are data collection and data upload when the device is synchronised. The handheld logs data on local flash memory. When the device is synchronised to a desktop, a Java applet on the desktop connects directly to a database and uploads data. Once successfully uploaded local archives can be deleted. Visualisation clients can then connect directly to the database. Figure 2 shows the system architecture in this case. Aside from the 3D visualisations described in Section 3, we have built simple 2D graphical user interfaces that can extract data based on time range and geographical region.

2.3 Collaboration Mode

In order to support collaboration between scientists in the field and in the lab, we are extending the infrastructure of the Equator City project [11]. The City project has been building collaborative systems on top of the EQUIP system [12]. The City project has built a number of collaborative systems which all support cross-media collaboration. That is, participants share a “virtual” space and use media resources as required by the situation rather than having a preset notion of the collaborative workspace. They might access map tools, shared audio, web pages and 3D graphical views. Some participants may be in a physical space that is modelled by the resources in the system. Other participants will be remote, relying on the contextual information supplied by, e.g. the map, to establish and maintain collaboration [13].

In recent systems work, the City project has been extending previous systems to the urban context. The current systems use GPS location and mapping tools to discover resources with which to effect interaction and collaboration. We plan to integrate maps of pollution as a visualisation layer within this framework. Figure 2 shows

how the different components integrate. For more details see [14].

3 DATA MODELLING

3.1 City Modelling

Because pollution can vary drastically based upon local building configuration, we decided to base visualisations on 3D models of the urban fabric. We used three principal resources: Ordnance Survey vector maps, LIDAR (Light Detection And Ranging) height data, and aerial photography. In addition we have started to integrate procedures for modelling facades of building.

Figure 3 shows example vector maps from the Land-Line data set. All maps are © Crown Copyright. Land-Line is supplied as tiled data, with each tile comprising 500m x 500m. These are usually distributed in Neutral Transfer Format (NTF), and we can use either this or the OpenGIS Consortium’s Geography Markup Language (GML). These vector maps sometimes contain errors in polygons such as missing edges. We build a complete Delaunay triangulation from the vertices and then classify the resulting triangles into various sets, corresponding to layers in the original vector map such as buildings, pavements, water and roads, see Figure 4.

Once the complete area has been classified, the buildings can be extruded to give them heights. This is easily done at this stage, because each vector line has been uniquely attached to two planar triangles, so edges of buildings can be duplicated, one edge raised and the façade polygons inserted. Figure 5 illustrates the result.

At this stage, the model comprises a planar floor. The Land-Line data contains spot heights, but these are too sparse to construct a smooth surface. A better data source, if it is available, is LIDAR. This gives spot heights at densities of typically 30 points per 100 square meters (50 times higher than Land-Line). The horizontal accuracy is 1.5m in the worst case due to uncertainties related to the attitude of the survey aircraft. The vertical accuracy is about +/- 15cm. This data can be used to give both the height of buildings in the previous extrusion step and to construct a terrain height for the ground. To create a smoother terrain, the inverse distance weighting interpolation described in the following section can be used.

Once the terrain is non-planar, a few new issues need to be addressed. The first is that individual buildings will not have horizontal edges. This is shown in Figure 5. Also, water features need special treatment, since they will rarely be planar due to surveying and interpolation



Figure 4 Constrained Delaunay Triangulation

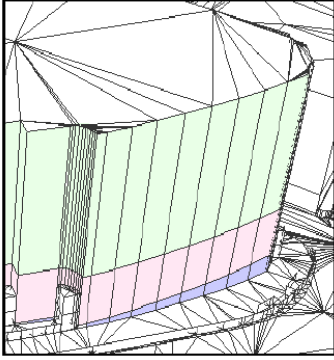


Figure 5 Triangles comprising a building façade

processes. These need to be flattened and appropriate banks put in where water features adjoin other features.

If aerial photography exists, then it can be draped over the mesh. We have used sections of the Cities Revealed data set from GeoInformation International for current demonstrations. Finally, other features can be added such as more complex facades, and street furniture.

3.2 Data Modelling

The input data for the pollution model is a stream made of a GPS position (x_i, y_i) and pollution data f_i (CO in parts per million). To make a 2D field representation, we first extract a temporal section of the data. The resulting data set is treated as a set of irregular scatter points.

One of the most commonly used techniques for interpolation of scatter points is inverse distance weighted (IDW) interpolation. Inverse distance weighted methods are based on the assumption that the interpolating surface should be influenced most by the nearby points and less by the more distant points. The interpolating surface is a

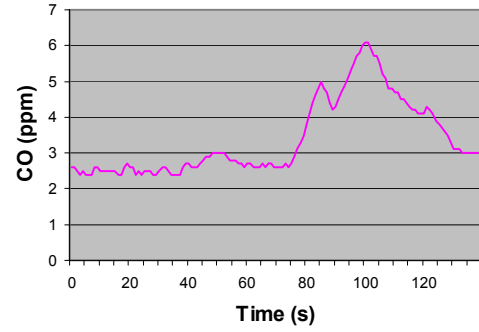


Figure 6 Raw data from a segment of a path near UCL

weighted average of the scatter points and the weight assigned to each scatter point diminishes as the distance from the interpolation point to the scatter point increases.

The simplest form of inverse distance weighted interpolation is Shepard's method [15]. The equation used to find the value at position (x,y) is:

$$F(x, y) = \sum_{i=1}^n w_i f_i$$

where n is the number of scatter points in the set, f_i are the prescribed function values at the scatter points (e.g. the pollution values), and w_i are the weight functions assigned to each scatter point. The weight function is:

$$w_i = \frac{h_i^{-p}}{\sum_{j=1}^n h_j^{-p}}$$

where p is a positive real number (typically, $p=2$) and h_i is the distance from the scatter point.

The effect of the weight function is that the surface interpolates each scatter point and is influenced most strongly between scatter points by the points closest to the point being interpolated.

4 RESULTS

4.1 Raw Pollution Data

The data shown in Figure 6 was collected on a path starting in UCL's front Quad, and walking up towards Euston Road. Before reaching Euston Road, the user crossed to the other side of the road, and the peak was reached when they were stood near the traffic lights at the junction of Euston Road and Gower St. The peak capture was 6.1 ppm.



Figure 7 Overview of area around St Paul's

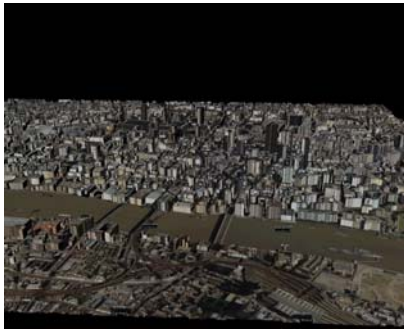


Figure 8 Oblique view of St Paul's area

4.2 City Models

The results of the city model generation are shown in Figures 7 and 8. These show a model of nine sq km around the St Paul's area of central London. The model comprises 1.2M polygons. Un-optimised this renders at 3-4 frames a second on a PC with GeForce4 graphics accelerator. An ongoing theme of research at UCL is interaction and interactive rendering of large-scale urban models [16].

4.3 Combined Data

Our aim in combining data is two fold: to support visualisation by placing the data in the context of the situation where it was gathered, and to support remote collaboration where one participant is using a virtual environment display to collaborate with a colleague in the field. Figure 9 shows views of the junction between Gower St and Euston Road. The blue line represents the recorded path from the GPS receiver. The inaccuracy of GPS location can be noted since the carrier walked along the centre of the pavements except when crossing Gower St.

In the visualisation in Figure 9 we present the pollution interpolation by colouring the roads. In order to maintain a high frame rate, we only interpolate the pollution level at each vertex of the road polygons using the inverse

distance weighted interpolation. We then use the built in Gouraud shading algorithm of standard graphics drivers to do a smooth interpolation. This typically uses a bi-linear interpolation. For a more accurate view, a 2D raster image can be calculated at some fixed spatial frequency. In Figure 9 the junction is obviously the most polluted area.

5 DISCUSSION

5.1 Flocks of Devices

At the time of writing we are deploying five sensors to build a denser and more extensive map of pollution. This poses several challenges in data modelling. We are looking at ways of establishing relative densities of different areas by removing gross temporal trends in the data such as the effects of weekday commuting.

The devices are being given to staff that commute to work by bike. Thus we will have a number of similar paths at similar times of the day. GPS location accuracy will probably be insufficient to establish variation across a road or from road edge to building side. We will look at filtering the GPS logs based on invalid intersections with buildings and variation between similar commuting routes. To establish differences between road and pavement pollution levels, we will be able to rely on different classes of device carrier. Logs from pedestrian users will allow us to see how pollution varies across a street, and how the experience of pollution varies for different users of the same street.

5.2 Grid Usage

A goal for the next couple of months is to integrate the pollution-monitoring device with the GRID network infrastructure provided by partners in the Equator e-science projects. This is described in detail in [9].

Since the pollution monitoring devices have no permanent connection, it is not possible to make them GRID services. It is also debatable whether it would be interesting to have a GRID infrastructure that supported the granularity of one service per device since each device is quite simple. Indeed we have argued that the interesting work is in the integration of data from multiple mobile sensors.

In data-logging mode (see Section 2.2 and Figure 2), it would make sense to wrap the synchronisation service, database services and 2D and 3D visualisation components as grid services.

Supporting collaboration mode is more difficult (see Section 2.3 and Figure 2). The elements that are shared with data-logging mode could be re-used, but the collaboration is supported by the Equip system. In

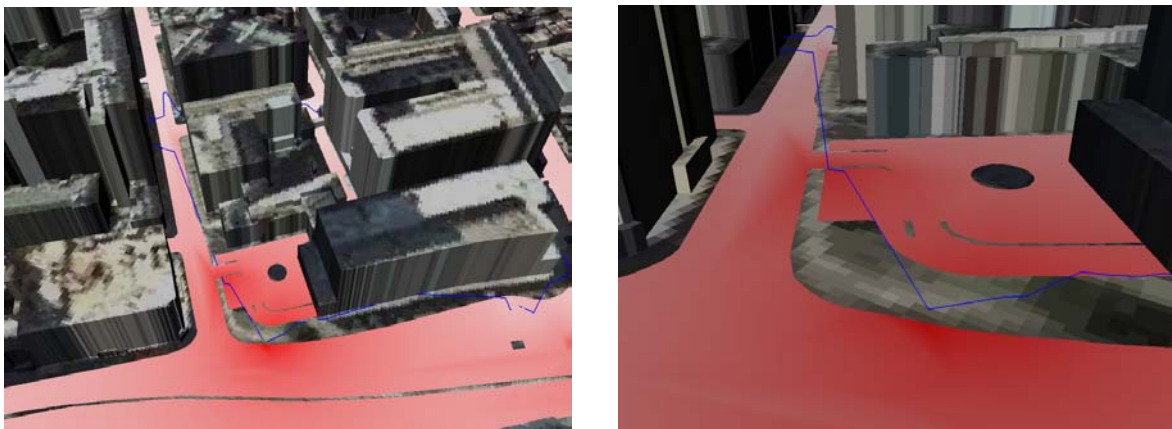


Figure 9 Views of the junction of Gower St and Euston Road

general, GRID tools such as the Access Grid [17] are targeted at high-end collaboration and not the mobile situation. They do not support a wide range of collaborative tools. We have suggested in Figure 2, that proxies could retrieve 2D and 3D models from the GRID-enabled visualisation services and publish them into Equip. The alternative, for each Equip client to have a GRID capability, is probably not possible given that some clients will be handheld computers.

6 CONCLUSIONS AND FUTURE WORK

We have demonstrated the feasibility of making dense maps of pollution using mobile sensing devices. This enables new types of monitoring that address local variation in pollution and also the levels of pollution that are experienced by different users of the urban space.

The next stage on device construction is to complete the two second-generation prototypes. Infrastructure work will proceed in two directions: integration of the various processes that support data capture, logging and visualisation with the GRID infrastructure being designed and built by the Equator e-Science projects; and real-time collaboration system where pollution data is shared in real-time amongst several participants. In the next couple of months, we expect to demonstrate a live connection between the UCL ReaCTor system, an immersive projection system supporting stereo graphics on four 3mx3m walls that surround the user, and a person out in the field using the monitoring device.

We hope to establish the infrastructure as a public infrastructure that can be shared or instantiated by other users wanting to investigate pollution. There is an opportunity to expand the system to many more pollution sensors, and perhaps even, make it a public understanding activity where people can build their own sensors and contribute data. This poses a significant challenge in verification and calibration of incoming data. As a first step, we are investigating integrating data from fixed

sensors into the model to act as calibration points for mobile sensors that pass the locations of the fixed sensors.

FURTHER DETAILS

This work is supported by EPSRC Grant GR/R81985/01 “Advanced Grid Interfaces for Environmental e-science in the Lab and in the Field” and the EQUATOR Interdisciplinary Research Collaboration (EPSRC Grant GR/N15986/01).

For a complete overview of the environmental e-science project, including the companion project on environmental monitoring in the Antarctic, see the EQUATOR website pages for this project [1]. For example data sets and more detailed specifications of the device see the web page [18]. We plan to make a public release of the software, and to host an example visualisation service at that address. For further information about the pollution-monitoring project please contact Anthony Steed (A.Steed@cs.ucl.ac.uk). For more details about the environmental e-science project in general, please contact Chris Greenhalgh (cmg@cs.nott.ac.uk).

REFERENCES

- [1] EQUATOR, “Environmental e-Science Project”, <http://www.equator.ac.uk/projects/environmental/index.htm> (verified 2003-08-13).
- [2] Environment Agency, “Air Quality – Carbon Monoxide”, http://www.environment-agency.gov.uk/yourenv/eff/air/222825/222913/?lang=_e (verified 2003-08-13).
- [3] “The Air Quality Archive”, <http://www.airquality.co.uk/> (verified 2003-08-13).
- [4] B. Croxford, A. Penn, B. Hillier, “Spatial Distribution of urban pollution: civilizing urban

- traffic”, Fifth Symposium on Highway and Urban Pollution, May 22-24, 1995.
- [5] Department for Environment, Food and Rural Affairs (Defra), “The Air Quality Strategy for England, Scotland, Wales and Northern Ireland”, 1999, available online at <http://www.defra.gov.uk/environment/consult/airquality/pdf/airstrat.pdf> (verified 2003-08-13).
- [6] Discovery Net, <http://ex.doc.ic.ac.uk/new/index.php>, (verified 2003-08-13)
- [7] “DAPPLE (Dispersion of Air Pollutants and their Penetration into the Local Environment)”, <http://www.dapple.org.uk/> (verified 2003-08-13).
- [8] Learian Design Ltd, <http://www.learian.co.uk> (verified 2003-08-13)
- [9] T. Rodden, C. Greenhalgh, D. DeRoure, A. Friday, L. Tarasenko, H. Muller et al., “Extending GT to Support Remote Medical Monitoring”, Proceedings of the UK e-Science All Hands Meeting 2003, Nottingham, Sept. 2-4.
- [10] S. Benford, N. Crout, J. Crowe, S. Egglestone, M. Foster, et al., “e-Science from the Antarctic to the GRID”, Proceedings of the UK e-Science All Hands Meeting 2003, Nottingham, Sept. 2-4. University of Nottingham, “The Equator Universal Platform”, <http://www.equator.ac.uk/technology/equip/index.htm> (verified 2003-08-13).
- [11] I. MacColl, D. Millard, C. Randell, A. Steed, B. Brown, et al., “Shared visiting in EQUATOR city”, Proceedings of the 4th international conference on collaborative virtual environments (CVE 2002), Bonn, Germany, p88-94. ACM Press, 2002.
- [12] University of Nottingham, “The Equator Universal Platform”, <http://www.equator.ac.uk/technology/equip/index.htm> (verified 2003-08-13).
- [13] B. Brown, I. MacColl, M. Chalmers, A. Galani, C. Randell, A. Steed, “Lessons from the lighthouse: Collaboration in a shared mixed reality system”, Proceedings of CHI 2003, pp. 577-584, ACM Press.
- [14] EQUATOR, “City Project”, <http://www.dcs.gla.ac.uk/scripts/global/equator/moin.cgi/> (verified 2003-08-13).
- [15] D. Shepard, “A two-dimensional interpolation function for irregularly-spaced data”, Proc. 23rd National Conference ACM, ACM, 517-524, 1968.
- [16] A. Steed, E. Frecon, D. Pemberton and G. Smith, “The London Travel Demonstrator”, Proceedings of the ACM Symposium on Virtual Reality Software and Technology, December 20-22nd 1999, pp. 50-57, ACM Press.
- [17] The Access Grid Project, <http://www.accessgrid.org/> (verified 2003-08-13).
- [18] “Advanced Grid Interfaces for Environmental e-Science: Urban Pollution”, <http://www.cs.ucl.ac.uk/research/vr/Projects/envesci/>, (verified 2003-08-13).