

PhD Showcase: Using OpenStreetMap to deliver location-based Environmental Information in Ireland

PhD Student: Blazej
Ciepluch

Dept. of Computer Science
National University of Ireland
Maynooth
Maynooth, Co. Kildare. Ireland
bciepluch@cs.nuim.ie

PhD Student: Ricky
Jacob

Dept. of Computer Science
National University of Ireland
Maynooth
Maynooth, Co. Kildare. Ireland
rjacob@cs.nuim.ie

PhD Supervisor: Peter
Mooney

Environmental Protection
Agency
Clonskeagh, Dublin 14,
Ireland.
p.mooney@epa.ie

PhD Supervisor: Adam
C. Winstanley

Dept. of Computer Science
National University of Ireland
Maynooth
Maynooth, Co. Kildare. Ireland
adamw@cs.nuim.ie

ABSTRACT

This paper outlines research work-in-progress on delivery of location-based services for environmental information in Ireland. A prototype web map service (WMS) is developed to deliver map-based environmental information using a specially customised version of the OpenStreetMap database. This WMS must deliver a location-based information package to the user: maps of the area that the user is viewing, key state-of-the-environment indicator information for that geographical area, and links to where the actual data and further environmental information can be obtained. This information package must be presented in a way that best matches the environmental preferences of the user. These preferences are derived from a set of 'user profiles' of potential users of the WMS. Software tools developed during this work to derive geospatial products from the OpenStreetMap database are described and some of our observations of working with OpenStreetMap are discussed. The paper closes with the likely directions for the continuation of this research.

Categories and Subject Descriptors

H.2.8 [Database Applications]: Spatial databases and GIS; H.3.5 [Online Information Services]: Data Sharing

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 2009 ACM The SIGSPATIAL Special, Volume 1, Number 3, November 2009 Copyright is held by the author/owner(s). ...\$10.00.

Keywords

Metadata, Tagging, Environmental, OpenStreetMap

1. INTRODUCTION

The importance of environmental scientific research data, information, and knowledge cannot be underestimated [19]. The Irish Environmental Protection Agency (EPA) currently manages the STRIVE Environmental Research programme. STRIVE will provide up to 100 million euro in funding for environmental research in Ireland between 2007 and 2013. STRIVE aims to "disseminate the findings of individual research projects and the overall programme to the widest possible audience in a coherent and timely manner". The EPA believes that this will "ensure synergies across the wider research agenda and earliest the possible dissemination of research outputs into the formulation of policy" [7]. A major problem in delivering environmental information to the public is the problem surrounding access to national level mapping, spatial data, and associated cartographical products. It is the case in Ireland that spatial data is very expensive to purchase for use. In some cases spatial resolution may be inconsistent or mapping may be out of date. OpenStreetMap (OSM) is a project designed to create and provide free spatial data. It is based on an open-source philosophy, and combines wiki-like user generated data, with free access, allowing users to create, edit, download, and use OSM data in their applications. The research question here is to investigate how suitable OSM would be as the base mapping layer upon which to build a national level WMS for the distribution of environmental data and information in Ireland. Extensive literature searches have shown us that there is little published work on OpenStreetMap. This paper will contribute to knowledge in this area by outlining the work-in-progress from two collaborative PhD projects which commenced in November 2008 (PhD1) and April 2009 (PhD2) respectively. Both PhDs focus on research related to the development of location-based services for the delivery of

environmental information in Ireland. One of the first major milestones in this research and development of a prototype web map service (WMS) which will deliver map-based environmental information using a specially customised version of the OpenStreetMap database with particular emphasis on quantitative and qualitative outputs from environmental research in Ireland which is funded by the EPA.

1.1 Overview of Research Work

As stated above this project investigates using OSM as the base layer for the distribution of environmental data and information in Ireland. The other spatial data component is the environmental data and information. The types of environmental data and information include: Air Quality monitoring stations, air quality levels at various locations, water quality for rivers, lakes, and streams, location of environmentally sensitive facilities such as Urban Waste Water Treatment Plants, IPPC facilities, etc. This WMS must deliver a location-based *information package* to the user. Ideally the users current location will be supplied automatically from a location-enabled mobile device or selected from a web-based map. Within the information package the WMS will return map images of the geographical area that the user has specified, display of key state-of-the-environment indicator information for that geographical area, and links to where the actual data and further information can be obtained. Figure 1 shows a flowchart of the proposed WMS based on the OpenStreetMap database. Firstly, the OpenStreetMap data for Ireland is loaded into a PostGIS database. This process is described in more detail in Section 2.2. Mapnik software [15] is then used to render cartographical output from the PostGIS database. Mapnik is strictly controlled by fine-grained style files which control the visual aspects of the map output. Map tiles are generated and made available for the WMS. The OpenLayers javascript mapping container is used to provide a user-interface. Output is delivered in a variety of formats: as web-based maps, marked up text-based information, and content for display on mobile devices such as mobile phones. Content delivery will be driven by development of profiles of potential users of the WMS. This profiling will identify categories of users by analysing the environmental perceptions and biases of the user base. It will be also necessary to undertake a detailed comparison of the two principles approaches to providing the WMS: (1) provision of a basemap with the option to overlay additional vector layers or (2) pre-generated map tiles which have the layers of environmental information embedded into the map tile images.

1.2 User Profiling

In an effort to understand the target user community this research has adopted a social constructionist approach which is based on the belief “that the environment has many meanings for different people in society” [4]. The meanings associated with the environment are multiple and conflicting, and that people come to sense the environment in diverse, multifaceted, and complex ways. Cantril and Senecah [5] argue that “our conception of the natural environment is framed by our experiences which are tightly bound to local settings”. Regardless of our own personal life experiences most participants in studies of environmental perception tend to respond very positively to environmental attributes such as “trees”, “water-bodies”, “trails” [24] Gobster and Westphal

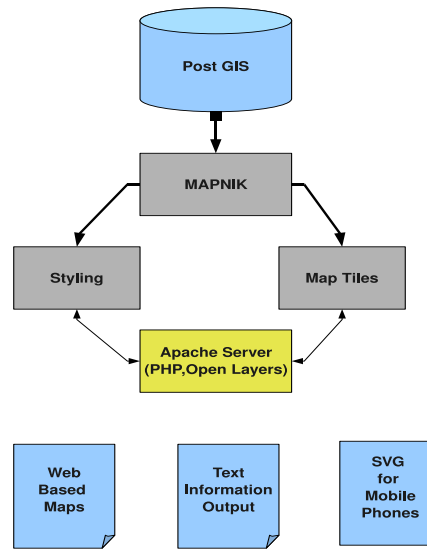


Figure 1: Flowchart for the proposed WMS based on OpenStreetMap data

[10] discuss the growing body of research that “confirms and elaborates upon the significance of nature in people’s aesthetic and recreational experiences and plays a crucial role in people’s physical and psychological health and well-being”. People’s scenic preferences for wildland, rural and urban landscapes show a strong favoritism toward natural features. Brody et al [1] describe empirical results which support the longstanding hypothesis that people living closer to a natural resource will be more familiar and concerned with its environmental quality. Brody et. al [2] show that physical position and proximity characteristics lend themselves to increased public perceptions of the potential negative impacts of climate change. Foster and McBeth [9] show that rural residents are more likely to be concerned with local environmental issues when they are defined in terms of quality of life features while urban residents are more focused on environmental protection issues. For example users who have lived in a rural setting all of their lives, users who grew up in an urban environment but now live in a rural setting, etc. To customise content delivery in this way it is necessary to undertake a study of how people perceive and evaluate their own natural environment. To build user profiles we analysed the search and download logs from the SAFER-Data website (<http://erc.epa.ie/safer>). This assumed that user preferences were embedded implicitly in the types of search words and combinations provided and the types of environmental data and information downloaded. This allowed us to identify three reasonably distinct categories of users:

- Local Users: Users interested in all environmental matters at a small-scale local level;
- Urban Users: Users only interested in the environmental matters at urban/city scale;
- Indication based-users: Those interested in gaining an overall picture of the *state of the environment* at a regional/national scale.

a Linux Cron task which schedules automatic download of the updated OSM XML file for Ireland every week and runs the necessary PostgreSQL statements to update our PostGIS database. During these updates it is not necessary to rebuild the PostGIS database. The `Osmosis` software can be used to compare the new `osm` file with the current PostGIS database contents and produce an OSM XML file containing only the data that has changed or being recently added in the past week since the last updated OpenStreetMap global QA/QC update. This allows a quick, automated, and easy updates to the base map of Ireland every week. The raw unprocessed OSM XML file contains all line, polygon, and point features. For each individual feature it includes every spatial attribute (called tags in OSM) associated with that feature. Only tags which validate against the OSM XML tagging schema are included. Figure 3 shows a flowchart of the processing applied to a raw OSM XML file when it has been downloaded. The PostGIS database in Figure 3 corresponds to the database which drives the WMS in Figure 1. As shown in Figure 3 the raw OSM XML data is loaded directly into the PostGIS database using the `osm2pgsql` tool [22]. The `Osmosis` tool is optional here as it can be used to support the loading process by offering functionality such as cropping. The `osm2pgsql` tool is a command line tool which takes an OSM XML file as input and loads this directly into PostGIS. The `ogr2ogr` [20] software package offers an extensive suite of spatial data manipulation functionality. There are several options for extraction of data from a PostGIS database and direct export to other GIS data formats. These formats include ESRI Shapefile, KML format, and MapInfo format. With the `ogr2ogr` tools one can execute SQL queries in the PostGIS database to perform query-based subsetting of the data within the database. This export functionality is particularly useful for data exchange with other similar or collaborative projects who may not have PostGIS installed or have software installed capable of accessing the OpenStreetMap data stored within our PostGIS database. The PHP software module, shown in Figure 3, is configured to write a GML representation of the OSM XML to disk. Separate GML files containing lines, polygons, and points are produced. For each feature a set of spatial attributes (tags within the OSM XML) are chosen. Conversion from the OSM XML coordinate reference system to Irish National Grid coordinates is also performed. These GML files are then passed to the `ogr2ogr` software which performs the file format conversion. In this project we exported predominantly to ESRI Shapefile format. The ESRI Shapefiles created in this way have been used for a related project in our department on Pedestrian Navigation using Location Based Systems. The `Update` module in Figure 3 refers to the insertion of additional spatial attributes for features in the PostGIS database. The `osm2pgsql` tool only extracts a default set of spatial attributes or tags for line, polygon, and point features from the OSM XML file. This is discussed in more detail in Section 3

3. EXTENDING THE OSM DATABASE

A disadvantage of both `Osmosis` and `osm2pgsql` is that only a specific set of spatial data attributes are loaded into the PostGIS database tables representing the OpenStreetMap for the region specified. It is necessary to access additional spatial attributes (tags) for application development. An example of this the requirement for our WMS to produce

```
<Rule>
<Filter>[highway] = 'unclassified'
or [highway] = 'residential'</Filter>
<MaxScaleDenominator>5000</MaxScaleDenominator>
<MinScaleDenominator>100</MinScaleDenominator>
<TextSymbolizer name="namega" face_name="DejaVu"
size="8" fill="#000" placement="line" />
</Rule>
```

Figure 4: Map rendering rule in the Mapnik style file to place the Irish name (namega) on all residential and unclassified highways

maps displaying native Irish language names on spatial features such as roads, streets, buildings, and other points of interest. In the OSM XML the `namega` tag is provided if an Irish language name is specified for the feature. The principal reason tags representing spatial data attributes such as native Irish placenames are not imported by these software packages is to allow map renderers (such as Mapnik) be developed for OpenStreetMap which automatically recognise the base set of tags imported. Without needing to work directly with the OpenStreetMap database developers can use these map renderers to quickly generate map output. For this project we require the availability of additional spatial data attributes for the lines, polygons, and points in the PostGIS database. We have written some software in PHP which takes the OSM XML file as input and automatically inserts the additional spatial data attributes we require into the PostGIS database. This is shown in Figure 3 where the PHP script updates the PostGIS database by automatically extracting the required tags from the raw OSM XML file. The rendering software Mapnik [16] then consumes a Mapnik style file (in XML format) and follows specific rules (see Figure 4) about which attributes are rendered on the map. This style file controls all of the cartographical parameters of the rendering engine: font control for map labeling, feature classification, icon placement, etc. For each additional tag we include in the PostGIS database we must specify a map rendering rule within the Mapnik style file. An example of a rule from a Mapnik style file is shown in Figure 4 where the residential roads are only rendered at a given scale with the text symbolizer specifying the style of the name label for the road feature on the map.

3.1 Examples of WMS Output

In this section we provide two examples of outputs from the WMS proposed in this paper. Stack-plume dispersions of pollutants from large chimney stacks follow Guassian air dispersion models [11]. The resulting calculations for air pollutant concentrations are often expressed as an air pollutant concentration contour map in order to show the spatial variation in contaminant levels over a wide area. In this way the contour lines can reveal the spatial relationship of air pollutants to areas of interest. Polygons illustrating the concentrations over the wider geographical area can be easily drawn as an additional layer. Such polygons can have an area for some local models of from 20 to 30KM from the stack location [17]. It is not an objective of the project to generate very accurate spatial stack-plume dispersion models. Rather our focus is to store the polygons representing the plume dispersions within the spatial database and sub-

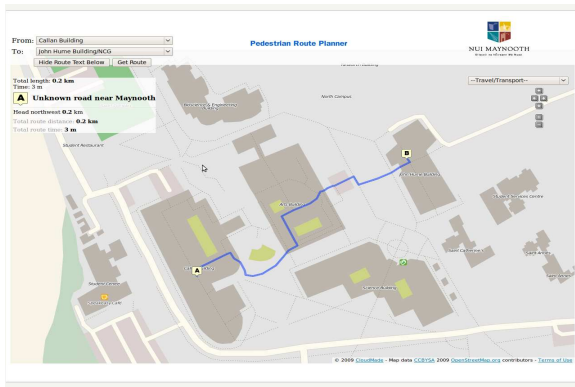


Figure 5: Using the Cloudmade API for shortest route calculations

sequently integrate these with other environmental quality data and information in the WMS output. The second example is provided by PhD2 which is investigating how to utilise the richness of the OpenStreetMap database in order to develop better context and location aware web services for mobile device users. OpenStreetMap benefits from the concept that the best knowledge about an area is possessed by people who live there or are familiar with the area [14][3]. Often such users can create maps containing a far richer set of spatial attributes and ground-based knowledge than maps extracted from aerial photographs regardless of resolution [13]. Figure 5 shows the CloudMade Web Lite API used to find the shortest walking path between two buildings on the NUI Maynooth campus. The routing algorithms within the API suggest a shortest route which advises a mixed route of external pavements and internal corridors inside buildings. PhD2 is currently investigating ways to add value to routing algorithms such as these by providing information on environmental features and information relevant to the spatial region in close proximity to the suggested route. Since proximity is a crucial concept in geographic relevance a major step is assessing which environmental features or information is *relevant* and *close-by*. This will involve establishment of appropriate distance functions to calculate the relevance of nearby environmental *objects*. Not all relevant features are easily contained within a uniform buffer zone around the suggested route.

4. CLOSING REMARKS

At the time of writing (June 2009) OpenStreetMap coverage in Ireland could best be described as poor. There is little or any coverage outside of the major conurbations. There are many examples where OSM within the major cities is poor. For example, in Figure 6 the region approximately 3KM east of Dublin International Airport is shown. This is a highly populated area yet the mapping in OpenStreetMap shows major sparse areas. We believe anomalies like this within the cities will be resolved as the usage of OpenStreetMap increases. Haklay and Weber [13] remark that even in London, where OSM started, large areas are undermapped and lack completeness. As of summer 2009 the number of users registered with OpenStreetMap who state their location as Dublin, Ireland is less than 50. This compares rather poorly with other European cities such as

London or Berlin. We would hope that as national level spatial data becomes more widely available and less constrained in terms of licenses then a better coverage of rural Ireland could be achieved. We looked at the OSM data for 4 European cities with similar population to Dublin city. Using a 2KM buffer centered at the locally recognised ‘city centre’ we counted the number of Points of Interest (POI) within the OSM dataset. The results are shown in Table 1 and are correct as of July 1st 2009. POI are probably the easiest geographical feature for OSM contributors to map and upload to OpenStreetMap. We can see from the table that Dublin has a relatively small number of POI compared to the other cities in our analysis. Haklay [12] performs an analysis of the *quality* of OSM data and shows that it is “fairly accurate” showing on average that OSM data is within about 6 meters of the positions recorded by the Ordnance Survey UK. Another aspect that Haklay finds affects coverage and quality is the number of OSM contributors in a specific geographical area with a direct correlation evident between a large number of contributors and better geographical coverage and spatial data quality. The OSM tagging schema (available

Table 1: OSM ‘Points of Interest’ in European Cities

City	POI
Dublin, Ireland	306
Sheffield, UK	591
Duisburg, Germany	449
Vilnius, Lithuania	252
Poznan, Poland	386

online from the OpenStreetMap website [21]) is increasingly being developed into a complex taxonomy of real world feature classes and objects. Haklay and Weber [13] state that “one of the most controversial aspects of OSM is that there are no limitations imposed on the tags users can attach to features”. However the authors conclude, and we strongly agree, that this flexibility benefits user application development because one can freely implement custom tagging schemas for specific end-user applications. We are currently investigating the feasibility of introducing special spatial attribute tags for representing environmental matters. The advantage, from a research point of view, of having local machine access to the entire OpenStreetMap database, is that we can easily introduce our own custom tags either by embedding within the original OSM XML file or by inserting new columns into the tables within the PostGIS database. The style-files for the map tile rendering tools are then updated to include tile rendering rules for each of the new tags. These tags include locations of instances of fly-tipping, locations of recycling facilities, and indicators of the quality of drinking or bathing water at water features.

Acknowledgments

Blazej Ciepluch is part of a five year project (2008 - 2013) called “Geoinformatics Services for Improved Access to Environmental Data and Information” led by Dr. Peter Mooney funded by the Environmental Protection Agency STRIVE research programme. Ricky Jacob is part of a project led by Dr. Adam Winstanley. Dr. Winstanley is the Principal Investigator for the Location Based Services (LBS) Strand of StratAG (Strategic Research in Advanced Geotechnologies). The assistance of Dr. Tracey Ni Mhaonaigh (Irish



Figure 6: An area of urban Dublin, Ireland which has no coverage in OpenStreetMap (June 2009)

Department NUI Maynooth) is acknowledged for providing us with street and road names in Irish for Maynooth.

5. REFERENCES

- [1] S. D. Brody, W. Highfield, and L. Alston. Does location matter?: Measuring environmental perceptions of creeks in two san antonio watersheds. *Environment and Behavior*, 36(2):229 – 250, 2004.
- [2] S. D. Brody, S. Zahran, A. Vedlitz, and H. Grover. Examining the relationship between physical vulnerability and public perceptions of global climate change in the united states. *Environment and Behaviour*, 40(-):72 – 96, 2007.
- [3] T. Bruny, H. Taylor, and M. Worboys. Levels of detail in descriptions and depictions of geographic space. *Spatial Cognition and Computation*, 29(7):227–266, 2007.
- [4] J. Bush, S. Moffatt, and C. E. Dunn. Keeping the public informed? public negotiation of air quality information. *Public Understanding of Science*, 10(-):213 – 229, 2001.
- [5] J. G. Cantrill and S. L. Senecah. Using the [’]sense of self-in-place’ construct in the context of environmental policy-making and landscape planning. *Environmental Science and Policy*, 4(4-5):185 – 203, 2001.
- [6] B. Ciepluch. Maynooth region, north kildare, ireland. openstreetmap rendering. OpenStreetMap hyperlink www.openstreetmap.org/?lat=53.38272&lon=-6.59455&zoom=15&layers=B000FTF- last checked June 25th 2009, 2009.
- [7] EPA. Science, technology, research and innovation for the environment (strive) - an environmental protection agency programme 2007-2013. Environmental Protection Agency, PO Box 3000, Johnstown Castle, Co. Wexford. Ireland, October 2007.
- [8] EPA. Integrated pollution prevention control (ippc) licensing. IPPC Guidance notes <http://www.epa.ie/whatwedo/licensing/ippc/> Checked June 09, 2009.
- [9] R. H. Foster and M. K. McBeth. Urban-rural influences in u.s. environmental and economic development policy. *Journal of Rural Studies*, 12(4):387 – 397, 1996.
- [10] P. H. Gobster and L. M. Westphal. The human dimensions of urban greenways: planning for recreation and related experiences. *Landscape and Urban Planning*, 68(2):147 – 165, 2004.
- [11] S. Godoy, P. Mores, A. S. Cruz, and N. Scenna. Assessment of impact distances for particulate matter dispersion: A stochastic approach. *Reliability Engineering and System Safety*, 4(16):10–20, 2009.
- [12] M. Haklay. How good is openstreetmap information? a comparative study of openstreetmap and ordnance survey datasets for london and the rest of england. Technical Report Currently under review at Environment and Planning B: Planning and Design, Department of Civil, Environmental and Geomatic Engineering, University College London, August 2008.
- [13] M. Haklay and P. Weber. Openstreetmap: User-generated street maps. *PERVASIVE computing*, pages 12–18, August 2008.
- [14] V. Kalakoski and P. Saariluoma. Taxi drivers’ exceptional memory of street names. *Memory and Cognition*, 29(4):634–638, 2001.
- [15] Mapnik. Mapnik: A free toolkit for developing mapping applications. Mapnik Online: <http://www.mapnik.org/>, April 2009.
- [16] Mapnik. Mapnik the open source map renderer. Software available online at <http://wiki.openstreetmap.org/wiki/Mapnik> - last checked June 2009, 2009.
- [17] C. Mensink, K. D. Ridder, F. Deutsch, F. Lefebvre, and K. V. de Vel. Examples of scale interactions in local, urban, and regional air quality modelling. *Atmospheric Research*, 89(4):351 – 357, 2008.
- [18] MK. Merkaartor: An openstreetmap mapping program, 2009.
- [19] P. Mooney and A. C. Winstanley. Managing environmental reseach data at epa. Technical report, Environmental Protection Agency, PO. Box 3000, Johnstown Castle, Wexford, Ireland, April 2009.
- [20] ogr2ogr. ogr2ogr a software package from the OGR Library to converts simple feature data between file formats. Available from www.gdal.org/ogr The OGR Simple Features Library is a C++ open source library, June 2009.
- [21] OSM. Openstreetmap: Map feature tags. Available Online: http://wiki.openstreetmap.org/wiki/Map_Features Checked June 2009, July 2009.
- [22] Osm2pgsql. A utility program to convert OpenStreetMap data into a format that can be loaded into PostgreSQL. Software available online at <http://wiki.openstreetmap.org/wiki/Osm2pgsql> - last checked June 2009, 2009.
- [23] osmosis. Osmosis: A command line java application for processing osm data. Available for download at <http://wiki.openstreetmap.org/wiki/Osmosis>, June 2009.
- [24] M. M. Wagner and P. H. Gobster. Interpreting landscape change: Measured biophysical change and surrounding social context. *Landscape and Urban Planning*, 81(1-2):67 – 80, 2007.