GEOTRAVELDIARY: TOWARDS ONLINE AUTOMATIC TRAVEL BEHAVIOUR DETECTION

A. Bolbol [a] *, T. Cheng [a], A. Paracha [b]

[a] Dept. of Geomatic Engineering, University College London, Gower Street, London, WC1E 6BT UK – (a.bolbol, tao.cheng)@ucl.ac.uk
[b] ParachaTech Limited, St John Street, London, EC1V 4PY UK - atif@parachatech.com

Commission IV, WG IV/5

KEY WORDS: Transport, Spatiotemporal, GIS, Online, GPS, User-generated, Interface, Visualization

ABSTRACT:

Travel diaries are considered to be a very important source of information that benefits major applications such as travel behaviour analysis. They have also been proven to be a burden on users to maintain and recall exact details, as well as being slow, expensive and time consuming. However, the rise of Geoweb 2.0, crowd sourcing and user-generated content is changing the way data is collected and shared. Nowadays, travellers could record their tracks and are able to upload them directly into the web to see them and even share them. The aim of this paper is to develop an online user-generated travel diary system that records users' trails. The online server-side application allows the users to visualize their trails and manually edit, move or remove points from their tracks on a map-based interface. Such a map-based interface enhances the retrieval of travel information from the users' long-term memory. The interface also allows them to add such information as metadata to different parts of their trip, such as the purpose of the trip and the transport mode. The application also overcomes travel diary disadvantages by being faster, cheaper and by providing users with incentives provoking higher response rates. This in turn will compile a large database of travel information within a large city (such as London in this case). Therefore, this will provide a better understanding of travel behaviour, and hence providing a platform for developing a realistic algorithm for learning travel metadata and hence deducing it. The application has just been launched at the time of publishing this work, and further work involves testing the application on a large scale and adding more functionality.

1. INTRODUCTION

As cities get bigger and more populated the need to study how people travel around increases. Travel behaviour studies help modelling tourist activity, understanding the impact of people's travel habits on the environment and measuring time expenditures and quality of life etc. However, the main problem is the collection of travel data.

Conventionally, the location data is collected by using positioning sensors such as GPS devices. Along with that, metadata like the purpose of the trip and transport mode (walk, bike, car etc) is collected by users using travel diaries during or after the journey (discussed in section 2). This has been proven to be a burden on users, and many other users do not recall exact details, as well as being slow, expensive and time consuming (discussed in section 3). Therefore, a cheaper, faster, and inviting alternative method is needed to overcome these disadvantages. There is also a need to understand how people record certain events in their long-term memory and how they best recall them (discussed in section 4). Moreover, there is also a need to understand what intrigues people to participate in the travel information collection in order to obtain better response rates (discussed in section 5).

As the whole world is going online, many long processes that used to be carried out very difficultly (sometimes manually) need to be brought to a digital, more dynamic and interactive environment via the web. Among these is the collection of travel information. In this study, we suggest that overcoming all of these disadvantages could be by achieved by having an interactive, fast, and online system that takes advantage of all the new Geoweb 2.0 features. We also suggest using incentives that invoke user-generated content to be produced along with all the best methods of retrieving information from the human long-term memory without creating a sense of privacy invasion of users. The application GeoTravelDiary.com is developed and presented in this work and is discussed in section 6. Conclusions and plans for future work are discussed afterwards in section 7.

2. TRANSPORT MODELLING

Previous research has attempted to generate mechanisms that are capable of efficiently modelling how people daily move around and commute. In turn, that has a great effect on tackling and forecasting major problems such as traffic congestion decrease of public transport usage and tourism management.

* Corresponding author.

2.1 Travel Information

In order to model people's behaviour certain elements need to be recorded along with the different locations they have been to and the routes they have taken. The tracks could be collected using mobile positioning sensor devices such as GPS, surveys or diaries. The other important information that needs to be collected with that could be:

- Different ***travel modes*** undertaken (walk, cycle, car, tube, rail, bus or ferry)
- ***Trip ends***: the location where the trip started and ended (home, work …etc.)
- ***Exact track*** taken on/off the existing networks (road, rail, tube, side-walks or rivers)

2.2 GPS Errors

The errors on produced maps were mostly generated due to the GPS positional accuracy errors, which causes lateral displacement in GPS tracks. This displacement causes misrepresentation of the actual route used on the transport network. It also causes problems like the wandering effect which is an illusion of excessive positioning wandering from a stationary device. Another problem caused by this is a cold start in recording position, which means that the device doesn't begin recording position data at the outset of a trip. That in turn leads to a wrong detection of trip ends.

2.3 Automatic Travel Detection

Previous attempts have tried to automate the travel information deduction by producing inference models. A lot of them do this by taking GPS data as inputs along with other user's information and producing travel mode and trip purpose information in return (Liao et al., 2007). These researches use a variety of techniques of machine learning such as neural networks, Hidden Markov Models, Labelling… etc. They also vary from supervised to unsupervised learning techniques.

In this work, our attempt is to provide a tool where users could label their own travel data (provide feedback) such as "where the transport mode was changed" or "where does the user's trip start and end". Users will also be able to do corrections to the collected data due to the GPS errors. In turn, this will provide a large database of labelled travel data which could be used to enhance the accuracy of the travel detection method.

3. TRAVEL DIARIES

Travel information collection has proven to be a very hard process. Many examples have proven to be expensive, burdensome, dangerous, slow, difficult to plan, come back with inaccurate information or have low response rates. The following is a summary of the previous attempts to conduct different studies to collect travel data from either recruited or volunteered respondents. Along with that we shed some light on the various problems highlighted by different research.

3.1 Surveys, Diaries & GPS-Aided Diaries

One of the most obvious examples is holding travel surveys either by telephone or face–to-face. This has proven to be a quite expensive, time consuming and sometimes dangerous to accomplish in some areas (Stopher & Metcalf, 1996).

Another trend was to engage people by maintaining their own travel diaries through the telephone, paper and pen, or computer-based (Stopher & Greaves, 2007). It was reported however that travel diary surveys have very low response rates because they were a burdensome task and some users omit certain travel information due to fear of their privacy invasion (Wolf et al., 2003).

A new trend emerged as to use GPS devices in conjunction with traditional surveys (Stopher, 2008). Some research even analysed GPS data in order to minimize trip under-reporting through improved survey methods (Bricka & Bhat, 2006).

3.2 GPS with Prompted Recall Surveys

A further step was to base the diary on GPS devices and subject the respondents to undertake prompted recall surveys. Attempts have been made to process users' GPS data and provide maps to users of their tracks, transport mode, and trip ends. They were then asked to visually verify and validate any identified errors on the map (Stopher, 2008).

Prompted recall surveys were considered the best means to obtain good accurate results, where respondents had the opportunity to see their tracks on a street map background which helped them identify errors and misinterpretations. However, among the main problems with prompted recall surveys are that they still achieved very low response rates (Bricka & Bhat, 2006). And if responded to, the main problem was the respondent's accuracy in providing the inputs because they weren't very involved in the map making process which is a specialist process.

3.3 Real Time Diaries

With the advance in the new emerging technologies in the market place, other research attempted to use real time mobile technologies to collect location data and an input respondent-entered data together in the same process. This theoretically is the most practical and time saving method that a travel diary could undertake.

On the other hand, many problems emerged due to real time usage, among these are the following:

1. **Privacy:** A sense of giving up one's privacy usually arises unless there is a significant amount of time between when the data is collected and when the respondents fill in the surveys (Stopher, 2008).
2. **User's Task Completion Time Gap:** The time lag between the instant the user starts on a different activity and the instant he/she records his/her status.
3. **User's Burden:** Users usually seem reluctant to record everything they do, because the real time activity takes some effort/time and occurs on several occasions along the day.
4. **Battery Requirement:** A common problem with real time GPS devices is having a short battery life.
5. **User Incentive:** Another very common problem that challenges lots of the previous research is the user's incentives. Some paid the users money (Stopher and Metcalf, 1996), while others tried to use special access panels that are set up with the purpose of conducting several different types of surveys on the same group of respondents by using incentives (Roorda & Miller, 2004). The rest mostly used volunteers that in many situations don't complete the full tasks (Stopher & Greaves, 2007).

In the application presented here, we try to overcome these shortcomings by tackling all the problems raised by the traditional travel diaries. The application takes the advantage of prompt recall surveys by providing an online map with their tracks requesting to provide feedback. It also is meant to be used in a desktop environment after the trips have been done in order to overcome the real time issues such as privacy, task completion time gap and the user burden. Users also have the incentive to participate by sharing and analysing their tracks (as discussed later in section 5.1). Long life GPS devices are also used to overcome the battery constraint (section 6.2).

4. HUMAN INFORMATION RETRIEVAL

The way that the human brain collects, manages, organizes, stores and retrieves information is quite a fascinating yet puzzling process. It is a process that happens almost unconsciously in our everyday life. Understanding such a phenomenon and being able to maximize the brain's ability to perform these tasks; could significantly enhance the quality and speed of collecting accurate travel information feedback from users. In this section we discuss principles of human cognitive psychology, and how to use it in favour of obtaining valuable feedback from users.

In a typical scenario of a travel diary, the different phases by which information is collected, stored and retrieved could be summarised in figure (1). The figure illustrates two main phases; firstly, the data collection phase where the GPS data is collected and the user (consciously or unconsciously) uses his working memory to remember the most recent roads he/she took, then only the encoded information flows into the long-term memory. Secondly, the information is evoked to be retrieved when the user is interrogated about which routes and modes he took on his journey.

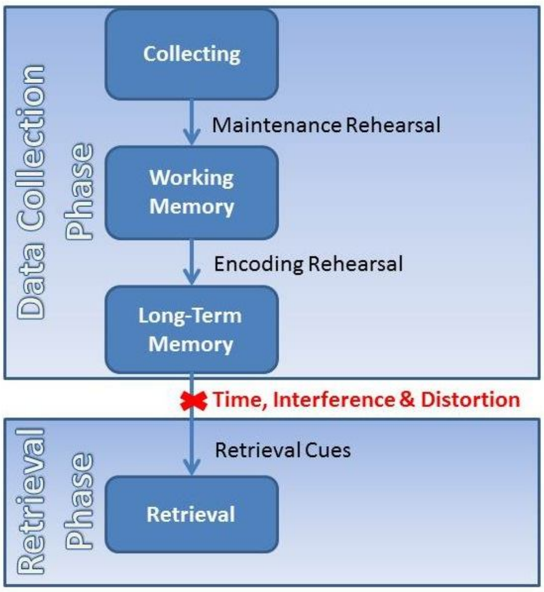

Figure 1. Information Flow in/out of the Human Brain

4.1 Data Collection - Information Encoding into the Long-Term Memory

The human brain attempts to rehearse information all the time. Psychologists distinguish between two types of rehearsal (Gray, 2001). ***Maintenance rehearsal*** is the process by which the brain holds information in the working memory for a period of time, and the ***encoding rehearsal*** is the process by which a person encodes this information into the long-term memory. Not all information that passes through the working memory is encoded into the long-term memory.

Only information that is elaborated, organised or visualised is encoded. ***Elaboration*** means that we remember things that capture our interest and stimulate our thought. The information could also be encoded by ***organisation*** by chunking the amount of items to be rehearsed into fewer items. The more the experience with the similar information the easier is the correlation between the different items. This is a similar to the case with master chess players being able to memorise locations of all pieces on the board (de Groot, 1965). Information also could be encoded by ***visualisation*** of items in mental images.

When people travel, they experience the route they are taking in different ways. A good objective here is trying to pass as much information as possible into people's long-term memory. That is because according to Gray (2001), information in people's long-term memory may exist without them being aware of it, and using certain retrieval cues later, this information could be evoked.

A natural advantage is that people usually elaborate on spatial navigation information because of the fact that they experience it in a real life context. People that take a similar route everyday (e.g. to work) tend to become experts of the area covered within the travel. This aids the objective of organising the route information and memorising connections between different streets. Also, much potential could be used to take advantage due to the fact that people tend to store mental images of their routes, by presenting users with photos of their different trip locations.

4.2 Causes of Human Information Loss

As illustrated in figure 1, there are some causes of memory loss such as time, interference or distortion. ***Time*** is a main factor memory loss where as more time passes the more the memory loses the stored information. Ebbinghaus (1885/1913) however, argues that most of the forgetting happens in the first short period then the rate of loss becomes increasingly gradual. Therefore, the sooner the users log in their feedback, the better, however, after a certain period (e.g. 2 days), the rate of forgetting stops increasing rapidly, and not a lot of information is lost onwards.

Another factor that contributes to forgetting is ***interference*** with one's memory. This relates to having to remember lots of other trips in the same area that users do. This is because the separately learned trips lose their distinctiveness and merge into one large pool (Gray, 2001). However, representing these trips on a map to the user, and having them segmented by time, loss of signal or by distance might overcome this problem.

The last factor causing memory loss is ***distortion*** caused by implanting false memories into users' long-term memory through leading or suggesting questions. This could be avoided by providing users with appropriate questionnaires.

4.3 Long-Term Memory Retrieval

Information is extremely organised inside our brains in the long-term memory like a dictionary (Gray, 2001). Perhaps the means by which we retrieve information from it (search within the dictionary) is the most fascinating part of the whole process. Mental associations are usually carried out in order to interrogate our long-term memory, like thinking about the colour "red" and remembering "strawberries", "apples", "London Buses" as a result. These are called ***retrieval cues***, and they help us probe our memory to find appropriate items to any category we think of.

Aristotle, more than 2000 years ago, proposed the law of association by contiguity, that is, if a person experiences two environmental events (stimuli) at the same time, or one after the other (contiguously), those events will be associated in the person's mind. And in turn, the thought of one will elicit the thought of the other (Horthersall, 1995). This could be useful in our specific scenario by presenting the users to view a virtual environment of their route, and that would elicit the thought of things like when they got off the bus, parked their car, or which turn they took. This is also enforced by another concept of retrieval called ***context-dependant memory***, where recall depends on similarity between the test environment and the original encoding environment.

4.4 Spatial Cognition

An important element to consider in this section is the spatial nature of information being dealt with. In order to account for different users' abilities to deal with space, we need to look closer into understanding different humans' spatial cognition. Spatial Cognition is concerned with the acquisition, organization, utilization, and revision of knowledge about spatial environments. These capabilities enable humans to manage basic and high-level cognitive tasks in everyday life.

Montello (2001) describes cognitive systems as including sensation and perception, thinking, imagery, memory, learning, language, reasoning, and problem solving. On the other hand, spatial properties include location, size, distance, direction, separation and connection, shape, pattern, and movement.

It is suggested that spatial knowledge of places is developed in a sequence of three stages:

1. **Landmark Knowledge:** Unique features that identify a place.
2. **Route Knowledge:** Travel routines that connect ordered sequence of landmarks.
3. **Survey Knowledge:** Corresponds to map-like mental knowledge. This could be enhanced by the exposure to maps of places, which is becoming very common nowadays with the availability of smart phones with many people (e.g. iPhone, Android… etc.).

People tend to store this spatial information as a cognitive (mental) map. This metaphor was introduced by Tolman in a 1948 paper to refer to a mental representation of spatial layout of the environment. The spatial information stored includes the landmarks, route connections, as well as non-spatial attributes (e.g. "the road where I rode my first bike"). Therfore, in the application presented here, we make use of Google Maps as the background for the interface which provides users with landmark locations along with their tracks. And even if users don't have good survey knowledge, they still can use the Street View background that provides photos of the route, which help invoking the retrieval cues to recall the exact route taken.

As stated in the previous section, there are many advantages of having GPS travel diaries with map-based recall surveys. A definite advantage is due to the fact that that when humans communicate space via language, they tend to expresses mostly non-quantitative or imprecise ('fuzzy') quantitative information. This is due to the lack of survey knowledge among most of people. And therefore, using map-based recall surveys are highly advantageous.

Another advantage with map-based recall surveys is that association by contiguity is achieved by presenting the user with the real map that evokes his/her mental map. As a result, association is achieved and hence retrieval cues evoke the memory to remember the user's state at different parts of the journey.

Other elements to account for are the similarities and differences between different groups. Some people are better at tasks such as way-finding, learning spatial layouts, or reading maps. There are many factors that may contribute to these variations in spatial cognition: body size, age, education, expertise, sex, social status, language, culture, and more. This creates uncertainty in the level of accuracy of the reported information. Therefore, the application builds a profile for each user compiling all their personal information to be able to draw associations.

5. GEOWEB 2.0

Nowadays, with the eagerness of people to take part in generating web content, concepts like crowd sourcing came into existence, which demonstrated the internet users' enthusiasm to contribute content (Howe, 2006). Web users feel they can express and post their version of n interest of more people about where everything is. This has matured people geographically and grew their location awareness.

5.1 GIS Users to GIS Technicians

Moreover, the availability of programming platforms like the Google maps and Google earth APIs that used very useful emerging technologies like AJAX, has lead to the concept of "Democratisation of GIS" (Goodchild, 2007). This also led more users to understand basic concepts of GIS, and even turned some into GIS specialists.

An obvious example is OpenStreetMap (OSM) where users contribute data for producing maps of the world online (Hacklay & Weber, 2008). This has encouraged and motivated people to collect their own GPS data and post it online. (Goodchild, 2008) argues that the reasons behind this motivation could be self promotion or the satisfaction that users derive from knowing that other people would be interested in seeing and even using their data. Other users are motivated by making information available to friend and relations.

In this work, we attempt to take advantage of this motivation and make use of it to replace the traditional travel diaries. We hope that this would pave the platform for bridging this link between the online Geoweb 2.0 capabilities and the previous horrendous process of collecting user transport data.

5.2 Online Applications

To the knowledge of the author, no application was found that serves this specific purpose. However, there are a number of similar applications that have different purposes and accordingly have different functionalities to serve that purpose. The mapmyfitness.com application enables users to upload the GPS tracks and states which type of sport they were undertaking. It returns analysis on their workout along with a representation of their tracks on a Google Maps background. maps.inersource.com also enables users to view their GPS tracks, routes and way points on different maps background for sharing and social interaction. Also, a similar interface is everytrail.com for sharing travelling experiences around the world on the web.
Very few applications provide the functionality of editing GPS tracks online from a map interface such as gpsvisualizer.com. Also, some desktop applications enable users to perform edit functionality (Viking software), but yet not having the functionality of adding any attributes to the data. In this work, we aim to have transport related edit functionalities to fill the missing gap in providing users with more transport-based GIS web functionalities.

6. GEO TRAVEL DIARY

The system developed in this study is an attempt to regulate and automate the travel diary concept in a Geoweb context. The system focuses on attaining the most adequate form of GPS data, in order to feed adequately into the travel activity inference model and obtain the best results. The system consists of multi-layered processes that flow in a waterfall manner from one process feeding into the following one.

6.1 Motivation - Aim of Application

After glancing through the literature, it becomes obvious that different phases of travel activity inference methods so far are unregulated. Also, there appears to be a lack of a systemized and guide-lined process of data collection for that purpose. Therefore, there appears to be an urging need for developing a system that is a natural step forward in the way travel diaries and follow-up surveys are collected and maintained. The following are the main aims of the application that make it an advantageous upgrade of the traditional travel diaries.

Online: A major advantage is having the system in an online environment which takes the burden off the users of saving their tracks and managing them.

Interoperability: Using a standard format of GPS files (GPX), which makes the input and output of the system flexible to read files from other GPS sources, and to be able to use the system's data on other platforms.

Visualization: Having Google Maps and Google Earth in a 2d and 3d environments respectively as a background map, enables users to have a better recall of their trip details. Another useful functionality in that context is using Google StreetView for viewing the trails, which simulates a realistic view of where users have been, which has in turn a major effect on their retrieval cues, where retrieval cues can provoke powerful recollections. This has led some researchers to speculate that there is a high possibility that all memories are permanent. The way to bring back these memories is by triggering some elements that were in the original environment of the original event when it took place. These elements are the retrieval cues, and having images (StreetView) simulating the actual route users took in the original event acts as some of these elements.

Trip Recall Time Gap: As mentioned in section 4, time is a major factor for memory loss. However, the rate of loss is very gradual after the first short period. Also, real time applications have numerous disadvantages like the ones mentioned in section 3. Among these is the problem resulting from users feeling a privacy invasion as a result of the real time notion. This trade-off is illustrated in figure 2. Therefore, this application gives the necessary time gap that the user needs to feel more at ease of the "reporting my location" notion. It also provides users with a good incentive by capturing their enthusiasm to upload their cycling, walking or driving tracks, which makes them eager to upload them regularly.

User Burden: The application also imposes fewer burdens on users by not having to input their metadata (or label their data) as opposed to mobile phone trackers.

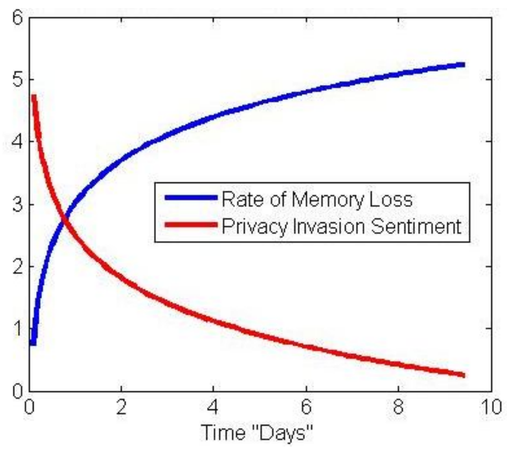

Figure 2. The trade-off between the Rate of human memory loss and the privacy invasion sentiment

User Interface: The interface also enables the users to change and input data visually from the map itself which makes the tasks required by the user easier, faster, and more accurate.

Incentives: As illustrated in section 3.3, participants need incentives in order not only to add data, but to spend time adding information to their collected data. The application therefore attempts to take advantage of the incentives behind user generated content the in Geoweb, as described in section 5.1. The interface allows users to run track analysis; so as cyclists & pedestrians enjoy watching their tracks and performances. It also gives feedback on the "green travel" ratio of the users' trips, which also generates a sense of awareness about people's travel habits and their impact on the environment. The interface also allows users such as holiday travellers to post their tracks in a shared community on the website adapting the idea of volunteered Geography (Goodchild, 2007) invoking self-promotion and sharing their experiences with friends.

6.2 Data Collection - GPS Devices

The application currently only supports data from GPS Devices of GPX format. The devices used in testing this application are developed by u-blox (see figure 3). These devices have a very reliable power saving property which enables the users to collect data for till more than 3 months. They also have accelerometers that switch the device on only if motion is detected, which increases the battery life and memory storage space very effectively. The devices are quite small in size too, which makes them attractive to use in tracking applications where they could fit unnoticeably in a jacket pocket or bag.

Figure 3. A typical u-blox GPS device used for this application

Other file formats from other devices are planned to be also accepted for upload in the future like Garmin GPX, KML. Users will also be able define their own XML formats in the future to be used in the type of files they upload. Also, plans in the future are in place to use formats from other devices that use different positioning sensors.

6.3 System Architecture

The conceptual design of the system is broken down into three phases/segments (as illustrated in figure 4). The first segment is the data collection where all tracks are collected whilst the users' trips using the GPS devices (or other sensors). The second segment is where the GPS processing happens using the ephemeris data. In the case of this test using the u-blox devices, the processing happens using a software package (YUMA) enabling a map interface to display the processed tracks. YUMA provides the output data in three formats: KMZ, CSV, and GPX.

The third segment is where the main user interaction happens in an online environment. We call the application the GeoTravelDiary (accessible from geotraveldiary.com). The application takes into account of all the specified advantages described in section 6.1.

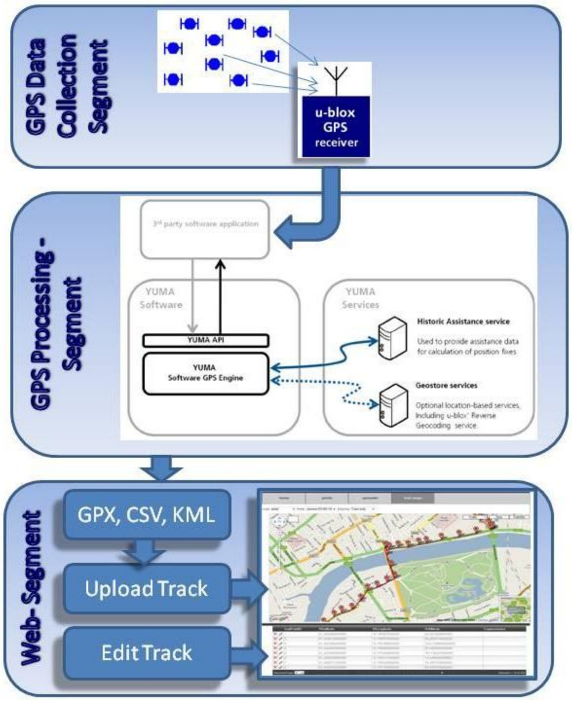

Figure 4. The system's data flow

6.4 GeoTravelDiary.com

6.4.1 Technology Used:

A.NET 3.5 framework is used along with a SQL Server 2008 Database, while the front end is developed in ASP.NET using XHTML, CSS, AJAX, JavaScript and JQuery

Object-oriented design is used and N-Tiered Architecture using Microsoft Enterprise Library 4.1 with static queries in stored procedures. Unlike normal web sites and applications, this site is developed using a post-back free/single page interface. Once the user logs in, they are taken to the Main page after which they remain on that one page till they log out. The menus on that page load new controls on the same page using AJAX, JavaScript and JQuery. This gives the user a much better response time to clicks and reducing network traffic by avoiding the transfer of the same content repeatedly. This sort of an interface, when required on the internet, is normally developed by JAVA embedding or using Flash/Silverlight or similar technologies.

Each one of them requires some sort of installation on the client side besides other issues like difficulty in integrating with web services and databases. Although the application has data driven content and menus, it has various hard coded features put in to speed up loading times. The application also arranges the storage of state information in a way to maximise efficiency.

The application is very scalable and the design allows rapid development. Free use of JavaScript is avoided and it is mostly done in tight bundles of pre-tested code packed together in controls. This allows the developer to build on the application and access JavaScript functionality using C#. This makes the application easy to debug during development and testing thus making upgrades quicker and cheaper.

6.4.2 Data Flow:

The data for the tracks is obtained by the users uploading their individual tracks on to the system. Once received, the data is parsed and stored as points, data relating to the points and relationships between the points on an online SQL Server database. Other useful information about the satellites is also saved in the RDBMS in appropriate tables.

We are able to run spatial queries directly on the database by using the special spatial features available in SQL Management Studio. The data is also retrievable into the ASP.NET application via Stored Procedures. In the application, the data is shown on its own, in grids, used to compute figures or to build graphs even as points and lines on maps. The data for the individual points and relationships between them, obtained from the database are used to display markers and to connect them together to show approximate tracks.

In the application, the data is used by grouping it together as objects to mimic the structure of the database. This allows easy manipulation of the data while still maintaining the integrity of the database. Since the data access logic is embedded into the lower tiers of the application, a front end developer can easily add more functionality to the application without having to worry about the complex data relationships.

6.4.3 UI-Interaction:

The main aim of the application is to mainly interact with the interface through the map. This enables taking more advantage of visualization. However, users could also do administrative tasks like changing their personal details through normal page interaction. Also, users will have to upload their tracks beforehand in CSV, KMZ, KML, or GPX formats.

The map background could be viewed in several flavours (Map, Satellite, Hybrid, Earth (3d), or Street View). Also, having the option of viewing Street View along with the Map enhances the user's experience to be able to link the tracks' location on the map with the images of the route taken.

The edit capabilities of the user could be classified into two types: Marker Edit or Data Edit. Marker Edits are edits to the Google markers on the map. These functionalities are such as dragging, deleting or adding the markers (see example in figure 6). Data Edits are edits done to the attributes of the points of within track. This includes editing information about the transport mode and type of location (work, home… etc.).

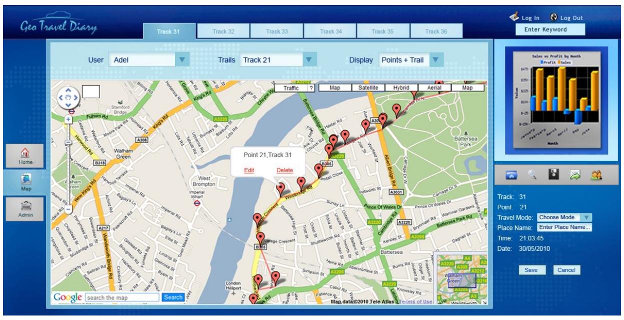

Figure 6. The application's interface showing point edit functions

Once a transport mode is selected for a point in the trip, all the points that follow are assigned the same mode until a new mode is selected for any point afterwards. Also, the colour of marker changes according to the transport mode, which gives an indicative visualization of when the users used different modes.

Users can also select to view their tracks by date, mode or user. Moreover, they can run analysis of how much percent of their travel was environmentally friendly or an average of how much calories they burned in a week, or even how many miles they drove their cars. These types of analysis are considered as an incentive to motivate users to accurately input their attribute information such as travel mode changes.

7. CONCLUSIONS & FURTHER WORK

Introducing such a system in an online environment tackling all the usability, response incentives, and visualization issues clearly has the potential to change the way travel diaries and travel information has been collected in the past. This has been shown where the application overcomes the main disadvantages of traditional travel diaries such as speed, cost, and response rates.

The introduction of a map-based interface where all the user edits are inputted makes non-GIS users perform GIS view and edit functionalities very easily. And the fact that it is in an online environment aids the objective of providing users with incentives such as web-contribution, emissions awareness and sharing experiences with other users. This also makes the process fast, easy to share, and useful for having a large database for running travel and traffic analysis. The application also provides the tools that provoke retrieval cues in the user's long-term memory by using background mapping along with Google StreetView tool which acts as a strong association tool.

Further work is to be done regarding sharing trips on various social networks (such as Facebook) to provide large scale awareness of transport issues and alternatives. Also, more functionality is to be added to the map interface in order to give the users more flexibility to add their information along with a better testing with a large number of users. Also, a second phase is to use the collected labelled data to develop an algorithm that learns and deduces the travel information automatically.

8. REFERENCES

Bricka, S.& Bhat, C., 2006. Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data, *Transportation Research* Record No. 1972, pp. 9-20.

De Groot, A.D., 1965. *Thought and choice in chess*. The Hague: Mouton.

Ebbinghaus, H., 1885; reprinted in 1913. *Memory: A contribution to experimental psychology*. (H.A. Ruger & C. E. Bussenius, Trans.). New York: Teachers College Press.

Goodchild, M.F., 2007. Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. *International Journal of Spatial Data Infrastructures Research*, vol. 2, pp. 24–32.

Gray, P., 2001. Psychology. 4th Edition. Worth publishers, New York, pp.126, 340-354.

Haklay, M. & Weber, P., 2008. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing*, 7(4).

Horthersall, D., 1995. *History of psychology*. 3rd Edition. New York: McGraw-Hill.

Howe, J., 2006. The Rise of Crowdsourcing. *Wired*. www.wired.com/wired/archive/14.06/crowds.html (accessed 26 April 2010)

Liao, L., Fox, D. & Kautz, H., 2007. Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields. *The International Journal of Robotics Research*.

Montello, D. R., 2001. Spatial Cognition. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*, pp. 14771-14775. Oxford: Pergamon Press.

Roorda, M.J. & Miller, E.J., 2004. Toronto Activity Panel Survey. A multi-instrument panel survey. *The 7th International Conference on Travel Survey Methods*. Costa Rica, August 1–6.

Stopher, P.R. & Greaves, S.P., 2007. Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*, 41(5), pp. 367–381.

Stopher, P.R. & Metcalf, H.M.A., 1996. Methods for Household Travel Surveys. *Synthesis of Highway Practice*, 236, NCHRP, TRB, Washington, D.C.

Stopher, P.R., 2008. Collecting and Processing Data from Mobile Technologies. *The 8th International Conference on Survey Methods in Transport*, Annecy, France.

Tolman, E. C., 1948. Cognitive maps in rats and men. *The Psychology Review*, 55 pp. 189/ 55, 189-208.

Wolf, J., Loechl, M., Thompson, M. & Arce, C., 2003. Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys, in P.R. Stopher and P.M. Jones (editors) *Transport Survey Quality and Innovation*, Pergamon Press, pp. 483-498.

9. AKNOWLEDGMENTS

We would like to express our gratitude to EPSRC, u-blox and KTN, the sponsoring bodies for this research, and Chris Marshall and Andy Yule of u-blox for their invaluable ideas and support. Also, a special thanks to all the colleagues that participated in the preliminary tests.