iLogLife

Final Year Project Functional Specification

Sarunas Vancevicius – 54455053

Supervisor: Prof Alan Smeaton

December 2007

0. Table of contents.

1. Introduction

1.1 Overview

The aim of this project is to see how accurately automatic summary of the day's activities can be generated from the use of logged GPS data and the surrounding Bluetooth devices in the local environment.

With the GPS logs for each day, it is possible to extract significant locations visited, while the logged Bluetooth device environment will give an insight of who was close by at each point during the day. Only limited GPS signal is available indoors, thus it is possible to know is the significant location indoors or outdoors by inferring the last known outdoor reading.

The collected data will be analyzed and a summary of day's activities will be generated on a blog-type web application. The user can just leave the key elements extracted from the data on the blog, or use those keywords as a stepping stone to manually write a more accurate description of the days events and expanding what the keywords initially provide.

The day's entry can include information such as:

Locations visited, both regular and significant, the arrival time and the departure time.
It will be possible to edit the location's details, such as the name given, more precise location/radius.
It might also be possible to class the significant locations to certain category/activity, for example, DCU would be classed as college, work place as work, certain friends' houses as social, etc.
It should also be possible to show the relevant Wikipedia articles relating to the location using www.GeoNames.org API.

Any "friends" that were detected while at that location.
This is possible because most mobile phones currently are Bluetooth-enabled. If Bluetooth visibility is set to on, the device should be detected if it is within 10 meters. Since Bluetooth devices have unique MAC address, it is possible to assign that device to a contact and assign a category; for example: family, co-worker, class mate, etc.

This data should be made searchable by location, friend, tag, and category.
For example searching for "dcu" will show all the visits to DCU.

This data should also be visualized on a graph, showing different locations visited during the day.

Such system has multiple uses; the main example would be as a productivity tool to write a blog of days' events, prompting the user with each day's keywords. Another use would be for people who have developed memory problems for example amnesia or early stages of Alzheimer's disease to remind of places visited and people met.

Because of the large amounts of data that will be collected, group formation and evolution, connection patterns could be also researched.
The user is required to carry a Nokia 770 Internet Tablet and a Bluetooth GPS device.

1.2 Glossary

GPS: Global positioning system. A system of satellites and receiving devices used to compute positions on the Earth .

Bluetooth: Bluetooth is a technical industry standard that facilitates communication between wireless devices such as mobile phones, PDAs and handheld computers.

NMEA 183: National Marine Electronics Association is a combined electrical and data specification for communication between marine electronic devices such as depth finders, navigation instruments, and GPS receivers.

Nokia Internet Tablet: is a small (141×79×19 mm) wireless handheld Internet Appliance running Linux OS.

Geocoding: Geocoding is the process of assigning geographic identifiers (e.g., codes or geographic coordinates expressed as latitude-longitude) to map features and other data records, such as street addresses.

Reverse Geocoding: Reverse geocoding is the process of returning an estimated area as it relates to a given coordinate.
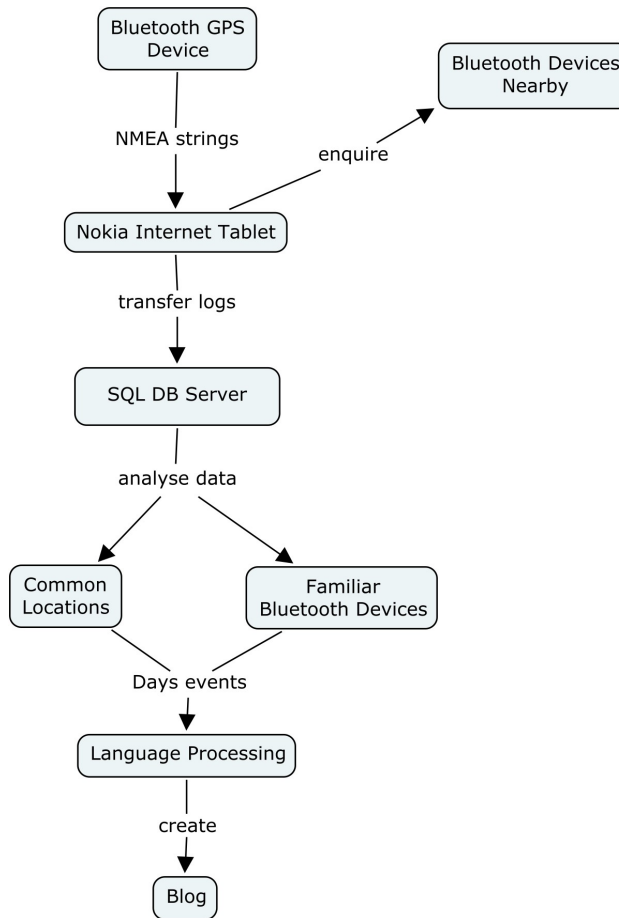
2. General Description

2.1 Product / System Functions

The basic functionality of the system:
>GPS data and Bluetooth Environment is logged on the Nokia Internet Tablet. This data then gets transferred to a database server at some point but it is not critical that this is done every day since the Tablet can store data on-board.
>This data is analyzed, extracting common locations and familiar Bluetooth devices
>Finally this data is presented on a blog-type application, in text and graphical format. The user can then write a more detailed post if s/he wants.

```
        ┌──────────────┐
        │ Bluetooth GPS│
        │    Device    │                    ┌──────────────────┐
        └──────────────┘                    │ Bluetooth Devices│
                                            │      Nearby      │
            NMEA strings                    └──────────────────┘
                        enquire
                │
        ┌──────────────────┐
        │Nokia Internet Tablet│
        └──────────────────┘

            transfer logs
                │
        ┌──────────────┐
        │ SQL DB Server│
        └──────────────┘

            analyse data

    ┌──────────┐         ┌──────────────────┐
    │ Common   │         │    Familiar      │
    │Locations │         │Bluetooth Devices │
    └──────────┘         └──────────────────┘

            Days events
                │
        ┌────────────────────┐
        │Language Processing │
        └────────────────────┘

            create
                │
            ┌──────┐
            │ Blog │
            └──────┘
```

2.2 User Characteristics and Objectives

This project could have 2 very different user types:

Teenagers and mid-20s people who are interested in social networks/blogging and they could use this as a memory aid and a productivity tool to expand their daily blogs and the people they meet. This could be expanded/integrated into a social network - by linking to a person's profile when clicking on their name/icon.

The system could be even expanded to just use a mobile phone with a GPS device or requiring the need to carry Bluetooth GPS device. The data plans on mobile phones are constantly falling (Vodafone now offers up to 50mb per day for 99c) so it would not be very expensive to upload the logged data to a central server. The users would be able to show were they are/where during the day, the friends they have met in those locations.

People who are at early stage of Alzheimer's decease, with the use of this system, could have an ability to review the places they have been and the people they could have met during the day, and as a result perhaps help stimulate their short term memory recall.

2.3 Operational Scenarios

A simple operation scenario for the system:
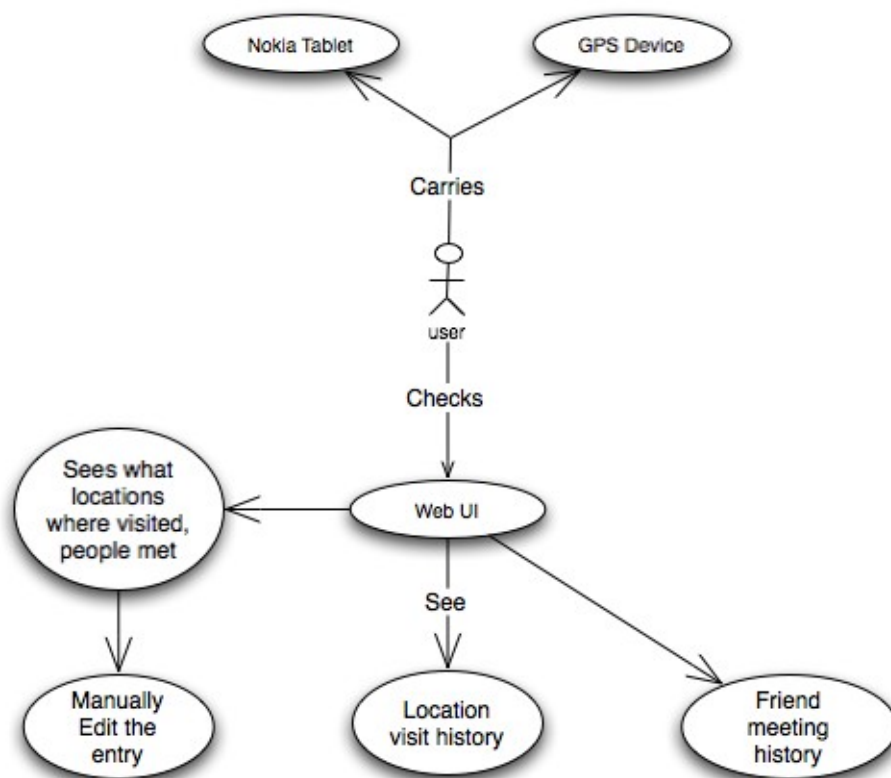
The user basically uses the web user interface to edit the automatic day summary generated entry and expands it manually. Clicking on any location will bring up the location visit history, same applies for a friend.

2.4 Constraints

Hardware constraints:

Nokia Tablet processing speed

Nokia Tablet log storage is on a flash device which can sustain only a limited number of write and erase cycles before failure

Battery life, Nokia Tablet has enough battery to log for a full day. The GPS device has a battery life of around 8 hours, thus requiring the user to charge it during the day.

Other constraints:

The user needs to carry this device at all times for the data logging to be done

The friend-detection relies on a people having their mobile phone Bluetooth setting to be set to visible. Also not everyone has a Bluetooth phone.

3 Functional Requirements

3.1 Logging of GPS data and Bluetooth Environment

Description

The GPS data and Bluetooth Environment must be logged on Nokia Tablet. The logged data will be stored in plain text or in an SQLite database.
Currently, the following data is logged:
Time, Date, Latitude, Longitude, Satellites in view (allows to know if we're indoors/outdoors), speed (direction of travel), Bluetooth device(s) MAC address and friendly name.

Criticality
This is essential for collecting the data

3.2 Transferring of GPS data to a server

Description
The collected data must be transferred to a more powerful central server eventually for further processing. But with a 2GB internal storage space, it is not urgent when/how often this data is transferred.

Criticality
This is essential. The Nokia Tablet does not have the processing power required to analyze the data.

Technical issues
The data must be transferred over a WiFi network. There are multiple issues with this:
> WiFi network connection might not be available, when to retry? What if the connection does not appear for a long time.
> Firewall might be blocking the access to the server where the data is stored

3.3 Analyzing the locations visited

Description
The location detection algorithm will either match a location to already existing location or as a new location.
A significant location will be defined as any area where significant time is spent.
The arrival time for a location will be defined as the first GPS coordinate within a location radius.
The departure time of a location will be defined as the last GPS coordinate which falls into the radius of the location.
Travel time will be defined as the time taken from the departure of location A to the arrival time at location B.

Criticality
This is an essential and critical part.

3.4 Analyzing of Bluetooth devices detected

Description
This part will do statistical analysis of the Bluetooth devices detected and classify them into:
Stranger, very few sightings of this device
Familiar Stranger, this device is seen often
User defined category, the user will have to manually classify this device as friend, co-worker, classmate, family, etc.
Then these devices will be shown that they were detected in the locations visited

Criticality
This is a critical part of the project which needs to be implemented.

3.5 Showing the location on the map

Description
Once a location is detected, it should be possible to see the location coordinates on Google Maps, with the adjustable radius for the location.

Criticality
This is a high priority function to help properly define a location, as a sub-locations will not be dealt with (instead of detecting that significant time was spent in the library or computing building within DCU, it will be just detected as DCU).

3.6 Getting the location name from geonames/local DCU gazetteer

Description
With the help of a reverse geocoding service from http://www.GeoNames.org/local and also a gazetteer available to this project from the CDVP, it should be possible to automatically identify the place name and use that for new locations.
The user can then change it to something more appropriate to be identified.
Criticality
Low priority function.

3.7 Display the analyzed days data in a Blog

Description
The analyzed data in human-readable format will need to be displayed in a Blog-style web application.
The user should have the functionality to manually correct the information presented.
This data should also be represented graphically. This could mean the user could associate certain images obtained from Flickr's geotagged photos to a certain location.
Another possibility is to acquire a photo for a friend from Facebook/Google Contacts API, or just manually upload a photo.

Criticality
High priority function.

3.8 Display closely related Wikipedia articles for the location

Description
The GeoNames.org service provides the ability to find nearby Wikipedia Entries which are Geotagged for given lat/long, and this could be integrated to provide automatic information for the location, to be incorporated into the automatically-generated blog.

Criticality
Low priority function

3.9 Integrate Google Calendar in the Blog entry

Description
Google Calendar provides an API to access events. It would be possible to display the events as part of day's entry.

Criticality
Technical issues
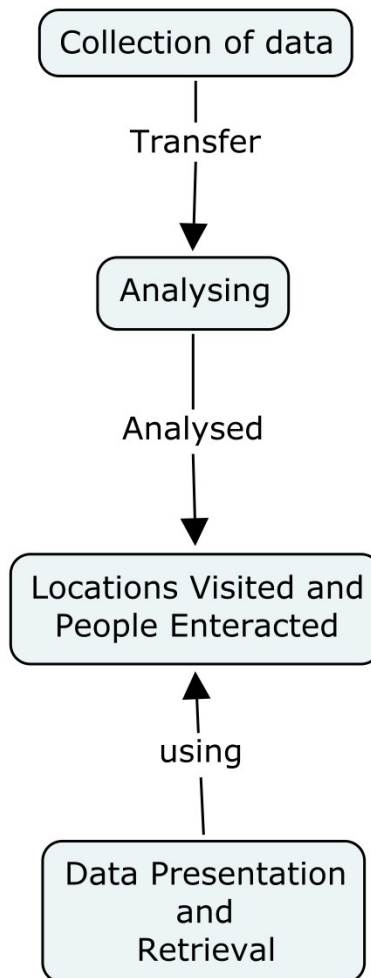Dependencies on other requirements

4. System Architecture

The system architecture has the following major components:

> Collection of data: This involves logging of the GPS data and the Bluetooth environment on a Nokia Tablet.
> Analyzing the collected data: Extracting significant locations and friends identified by Bluetooth devices.
> Presenting the user with the analyzed data and let the user edit if desired.

```
     ┌──────────────────────┐
     │  Collection of data  │
     └──────────────────────┘
                │
            Transfer
                ↓
        ┌─────────────┐
        │  Analysing  │
        └─────────────┘
                │
            Analysed
                ↓
   ┌────────────────────────┐
   │  Locations Visited and │
   │   People Enteracted    │
   └────────────────────────┘
                ↑
             using
                │
     ┌──────────────────────┐
     │  Data Presentation   │
     │         and          │
     │      Retrieval       │
     └──────────────────────┘
```

5. High-Level Design

This section sets out the high-level design of the system. The following diagram outlines the high-level design:
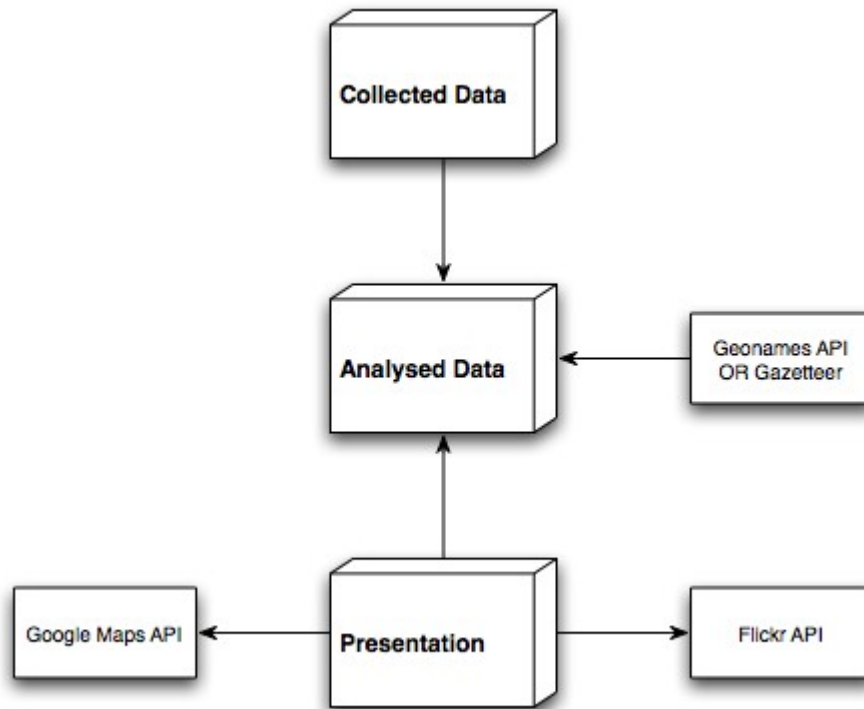
The collected data is gathered by the Nokia Tablet. This data is then analyzed and Geonames/Gazetteer is used to obtain location names. Finally, the User Interface presents the analyzed data and also uses Google Maps API and Flickr API to present the data visually (location, photos of the location).

6. Preliminary Schedule

Outlined schedule:
> Now - 4 Feb: Research and start gathering GPS and Bluetooth neighborhood logs.
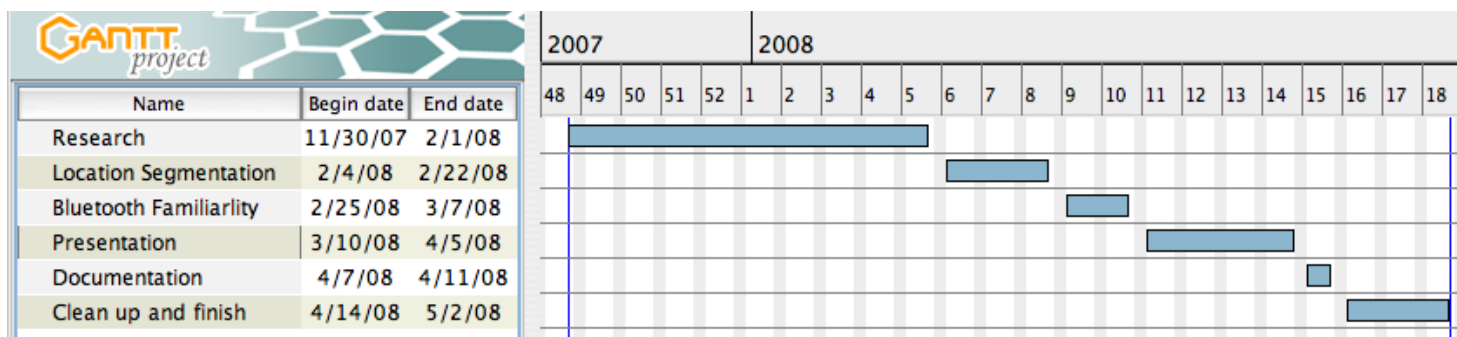> 3 weeks: Location segmentation
> 2 weeks: Bluetooth familiarity
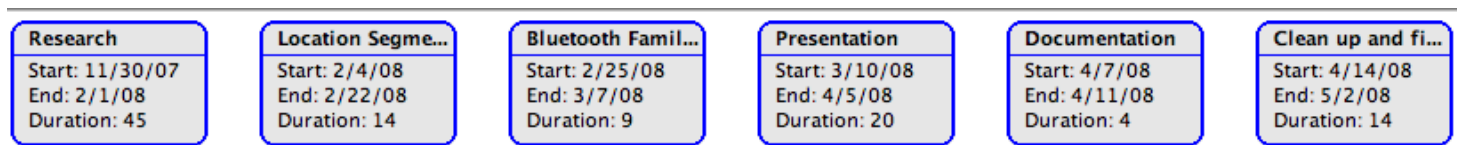> 4 weeks: Interface design and presentation
> 1 week: Documentation
> Remaining time: Cleaning up and finishing everything off

6.1 Gantt Chart

6.2 Pert Chart

7. Project Management

7.1 Version Control System

Subversion (an open source application for revision control) was chosen as the project version control system.

7.2 Documentation

To do an automatic project code documentation there are few options:
>   Doxygen, mature documentation tool.
>   pydoc: Python based documentation generator.
>   http://lfw.org/python/pydoc.html

7.3 Project Blog

The progress of this project will be updated on the project blog on: http://iloglife.blogspot.com/

7.3 Choosing programming language

I have chosen Python due to the fact that it is a high level programming languages and there is a broad range of modules available which means that things can be done very quickly providing efficient use of the programmer's time.

8. Appendices

8.1 Similar projects

My Life Bits, Microsoft: http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx
Reality Mining, MIT:http://reality.media.mit.edu/
Contect Phone, University of Helsinki, http://www.cs.helsinki.fi/group/context/
Jaiku, http://www.jaiku.com/