

6CCS3PRJ Final Year LOST

Final Project Report

Author: Joshua Simpson

Supervisor: Andrew Coles

Student ID: 1225043

April 16, 2015

Abstract

Navigational tools have been developing rapidly over the past couple of decades - at the moment, capable of providing us with a location within a few meters of where we are. Unfortunately, they still have not progressed to a point where a user can use location-based services from within a building to an accurate degree - an advance that would open new avenues of development in contextually aware services

This report showcases an extension of works that use wireless access point data to provide contextually aware location based services at a building - the end result in this being a mobile application package providing location information and route-planning functionality, whilst also crowd-sourcing contextual data. The main aim of this entire package being to demonstrate the value of contextually aware systems to different types of users.

Originality Avowal

I verify that I am the sole author of this report, except where explicitly stated to the contrary.
I grant the right to King's College London to make paper and electronic copies of the submitted work for purposes of marking, plagiarism detection and archival, and to upload a copy of the work to Turnitin or another trusted plagiarism detection service. I confirm this report does not exceed 25,000 words.

Joshua Simpson

April 16, 2015

Acknowledgements

Dr Andrew Coles - for his patience, expertise, and lulzy memes

The Hackathon Community - for their amazing support through difficult times

Contents

1	Introduction	2
1.1	Project Aims	3
1.2	Report Structure	3
2	Project Background and Main Points to Consider	4
2.1	Location Services	4
2.2	Context-Aware Services	6
2.3	Crowd Sourcing	7
2.4	Platform	8
2.5	Ethical Concerns	9
3	Requirements	11
3.1	LOST - Android Application	11
3.2	Backend	12
3.3	FOUND - Android Application	13
4	Design	14
4.1	Component Diagram	15
4.2	LOST	16
	Bibliography	18
A	Extra Information	19
A.1	Use Case Diagram	19
A.2	Wireframes	20

Chapter 1

Introduction

Location based services have been developing over the past few decades, and are prevalent in several aspects of every day life - location based services range from the route-planning software that is used to figure out a morning commute, to live time-tabling services that use bus or train positions to give accurate estimates. It is safe to say that in some way or another, location services are prevalent in modern day life.

However, GPS¹ services only work to a certain level; specifically, they only work to within a few feet at street level[5]. Whilst this is still a feat, it leaves a large market gap for contextually aware services inside large buildings.

Contextually aware services are another rapidly developing technology, providing users with information or activities specific to their current location or situation, determining the suitability for such events by using sensor data².

There are previous studies on using local area wireless access points to provide this sensor data at a level more precise than GPS can currently allow, but they are usually run in controlled environments, deriving a single mathematical equation in order to determine distance from the access point with high precision[4]. Unfortunately, this is impractical to apply as general use outside of a test environment.

¹ Global Positioning System - a system using satellites to provide precise location information

² Data such as location, temperature or time

1.1 Project Aims

The main goal of this project is to provide a proof-of-concept mobile application that can provide location and route-planning services whilst also delivering contextually aware services in a specific building - all without using GPS - demonstrating the possibilities that this area of application development holds for developers. Around this main feature, the project will aim to create software suite allowing for the efficient setup of a 'contextually aware building', as well as provide practical functionality to a variety of user groups (by using the application to crowdsource data), to demonstrate the potential in this field. To achieve this, I will be collecting data on user location and wifi failure³ throughout the building and providing a backend that displays this data in a useful manner.

To demonstrate a successful project, a large building with many potential locations is useful, as it shows the accuracy of the location-finding and route-planning algorithms. To this end, the building that I am using to test this application is King's College London's Strand Campus⁴. With roughly 17,000m² of space and 9 floors to utilise, the application would not only be thoroughly tested but also most likely welcomed by King's 25,000 students[8].

1.2 Report Structure

This report will begin with a background on the project - taking a look at location based services and contextually aware services and the applications of both that are already available, along with an analysis of projects or papers that have attempted to solve the problem of location services at building level. This data will then be collated to analyse key problem areas during implementation.

Following this there will be an outline of the requirements and specification for the project - showing both the project itself and the chosen extensions - complete with justifications on decisions made in the specification.

³A common occurrence in the chosen building

⁴ It has not escaped my attention that this will also help all the new first years find Waterfront

Chapter 2

Project Background and Main Points to Consider

2.1 Location Services

Defined as *"services that integrate a mobile device's location or position with other information so as to provide added value to a user"*[13], the origins of location information services date back to 1973, when the US Department of Defence developed GPS to overcome the limitations imposed by navigation systems in use at the time[10]. This original system has since been developed into an integral part of many mobile applications, including mapping and route planning applications, as well as forming the 'sensor' in a lot of contextually aware applications. There is a well established method that is used in order to ascertain location, called trilateration¹. Triangulation² is another method used to determine location, but is mostly inapplicable in the context of mobile devices³. Whilst I will not be developing an application that directly uses existing location technologies such as GPS, it is important to consider the practices used in such technologies when developing this project.

2.1.1 Application: Google Maps

Google Maps is easily the most widely used mapping tool on the planet, with over 54% of smartphone owners using the app on a regular basis in 2013 (for comparison, this was 10% more than Facebook at the time)[11]. Google Maps are a chief provider of location information

¹The process of finding a location by measuring distances, using geometry

²The process of finding a location by estimating the direction in which signals are coming from

³It is also commonly mistaken for trilateration

services, with advanced route planning and street-view functionality. They serve this data through their own API⁴. The most interesting aspect of their service (in relation to this paper, at least) is their recent addition of the use of Wireless Access Points for their location services. Whilst they do not rely solely on it for location, they do use the data to increase accuracy and speed through their Maps application[6]. Interestingly, Google collect this WiFi data by crowdsourcing through an 'opt out' scheme in the Google Maps application[14] - this may not be possible in the case of this project (as Google have used WiFi in a more assistive manner), but it provides a good basis from which to approach the problem.

2.1.2 Paper: Indoor Location Using Trilateration Characteristics

[5]

This paper is especially relevant to the problem, considering the use of Wireless Local Area Networks in order to achieve trilateration. This is achieved by measuring the signal strength for at least 3 access points, and using the signal strength to determine the distance from the access point. Once you have three accurate distances you can essentially create a 'Venn Diagram' of where the user is located.

The paper goes on to state that there are inherent inaccuracies with this positioning technique, due to the variable nature of radio signals[5].

To circumvent these inaccuracies, an elegant averaging method was developed, which started returning an accurate 'average signal strength' after 40 readings.

Issues with the trilateration approach

- As stated above, the trilateration approach has difficulties providing accuracies:

"Radio signals are extremely variable, particularly indoors, due to being reflected by obstacles or refracted round corners, known as multipath reflection. Environmental changes can also affect the signals, such as the number of people around. This means that the positioning technique is inherently inaccurate, with positions from raw Wi-Fi signals being in excess of 10m out. "[5]

- Whilst there is a solution designed to solve this problem in the case of the test environment, an application in a building such as Kings would require a number of different formulae in order to provide accurate signal strength averaging for different areas of King's⁵

⁴<https://developers.google.com/maps/>

⁵ Because of the variables such as distance between routers, building material, and the number of users in any area at a given time

- Finally, in cases where the averaging solution is easily applicable, there *should* be concerns regarding the impact on a mobile device's battery life when performing repeated scans and operations on the results of those scans.

2.1.3 Paper: More Stuff on Location - Fuzzy Location?

2.2 Context-Aware Services

Contextually aware services are services that use data that pertains to the user or application's current situation to provide actions or functionality specific to that scenario. It is defined by Abowd as

"any information that can be used to characterize the situation of entities (i.e., whether a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves." [3]

Contextually aware services use 'sensors' to fetch this information - examples include location data, temperature, even the camera on your phone could be used as a sensor for contextual services⁶.

2.2.1 Estimote

A rapidly accelerating tech start up, Estimote provides 'beacons' using low-power Bluetooth in order to send an ID⁷ to an Estimote enabled application - these can then be programmatically associated with locations, events, etc. and calculates its distance from the phone using RSSI (the received signal strength) inside it's own API, meaning that it can be used very effectively to provide location-based services with minimal effort. This is so far the the most prominent product in the market using the iBeacon base.

Currently the major drawback of the Estimote is the price, which limits Estimote beacons / stickers to small buildings - covering an area such as King's would be incredibly costly⁸, and any structural changes regarding the beacons would require significant effort to represent in an application developed to use Estimote. Another issue is that - as stated above - using RSSI to fetch location information can become very complex in a large building with lots of people in it (which would cause signal attenuation at different levels throughout the day).

⁶Consider how recent smartphones have software that stops the screen coming out of standby if it thinks it is in your pocket

⁷<http://estimote.com/api/getting-started/intro-to-beacons.html>

⁸<http://estimote.com/#jump-to-products>

2.2.2 Google Maps

As part of their vast toolset, Google Maps also has context-sensitive functions in its API - an example is their 'Explore' feature⁹, which will show destinations of a certain type that are in the user's local area. The Explore function even takes into account time of day and weather in order to provide the best possible locations to the user.

An interesting point to note is that the 'Explore' function was predated by a 'Search Nearby' function that was removed¹⁰ - possibly due to privacy concerns¹¹.

2.2.3 Paper: An Architecture for Data Collection and Processing in Context-Aware Applications

[12]

In this paper, Dario Salvi *et al* pose that whilst Context Aware systems are showing and developing a greater use in general applications, that the field is not "applied in the real world"[12]. The paper then goes on to propose an architecture that access and uses contextual data differently to provide the developer with more possibility to develop what is referred to as 'real world applications'.

Afterwards, the writers go on to set out three classes for collecting data - AtomicData¹², ContextData¹³, and Relationships¹⁴[12]. These classes provide an interesting base for the data collection part of this project¹⁵.

Overall the paper presents some very useful points on using contextual data to return usable information.

2.3 Crowd Sourcing

Crowd sourcing is generally defined as using a group of people or community to generate a needed resource. It is mostly thought of in terms of popular crowd-funding sites such as Kickstarter, or Indiegogo, but the term extends to other resources, such as data or ideas¹⁶. Generally, data crowdsourcing follows a pattern of clients paying for information that they

⁹<http://google-latlong.blogspot.co.uk/2014/07/spend-more-time-exploring-with-google.html>

¹⁰<http://www.jsonline.com/blogs/news/247119811.html>

¹¹Covered later in Ethical Concerns

¹²Singular pieces of information

¹³Compositions of more than one piece of AtomicData

¹⁴Links between two ContextData

¹⁵ContextData holds the most interest, as I will need to collect multiple pieces of data

¹⁶For example, Cracked.com - a satire website - uses its own staff, but also looks for stories contributed by its readers

need, and users that receive rewards to collect or generate this information.

2.3.1 Captcha

Captcha was originally invented as an identification system to help discern humans from robots when performing tasks online. This started out as just a verification system (which is growing more and more elaborate as robots evolve), but was eventually evolved into an aid to digitizing books¹⁷. The most interesting point about this is that most of the users of Captcha have no idea that they are crowd-sourcing a huge effort every time they solve a Captcha.

2.3.2 The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities

[9]

This article addresses a major problem with crowdsourcing data which Nielsen refers to as participation inequality, stating that user participation often follows a 90-9-1 rule, with 90% of users being "lurkers or users that will observe but not contribute"[9] . Nielsen also covers how to overcome this participation inequality (after briefly summarizing with "you can't"); one method in particular is of interest when considering the aims of this project and its associated application: 'make participation a side effect' - essentially the act of collecting data without the user having to do anything - this is less intrusive to the user's day-to-day use of the application, but can also have obvious ethical ramifications if the user is unaware if they are sourcing data.

2.4 Platform

Because of the nature of the application being developed, it makes the most sense to develop for a mobile device - whilst devices like laptops offer more power to process information, the dataset that needs to be taken does not require the extra resources. Alongside this, mobile devices lend themselves better to the data-collection aspect of the project, with background services that can perform operations whilst the phone is inactive in a user's pocket. Finally, it is easier for a user to take out their phone in order to ascertain their location than it would be for a user to take out their laptop to perform the same operation.

¹⁷<http://www.npr.org/2013/10/04/191620023/can-you-crowdsource-without-even-knowing-it>

With this in mind, the choice of platform then comes down to which type of mobile device to develop for - Android devices, or iOS devices. Whilst I could develop for both device types simultaneously using a platform such as Phonegap¹⁸, I have concerns about not using the native SDK for a platform - the abstraction provided by platforms such as Phonegap may reduce or restrict functionality that is necessary during development¹⁹. In the end, I decided to develop this for the Android platform alone.

2.4.1 Reasons for Android as a platform

The first point I considered whilst choosing a platform to use is the popularity of a platform - it makes sense to make the application available to the widest audience possible. In this case, Android dominates the market share, with 76.6% of smartphones using the Android OS, compared to 19.7% using iOS in the last quarter of 2014[7].

Another major concern when deciding upon device was the functionality provided by SDKs for both Android and iOS. When I reviewed the functionality for both devices, I was unsurprised to find that iOS does not allow access to lower-level functionality. What did surprise me is that statistics like wireless signal strength was amongst these lower level functions[2][1]²⁰. This restriction of access was a major concern²¹

2.5 Ethical Concerns

When developing an application that uses any form of personal information, there are definite privacy concerns. This is exacerbated when considering that there will be location information being returned to a server that is accessible by staff members. In this respect, it is important to not only adhere strictly to Data Protection laws, but also to consider the user's privacy and wellbeing. To this end, I need to make sure to collect the least amount of data I can, and to anonymise this data as much as possible in order to avoid sending any user details that would affect their privacy. Alongside this it would be prudent to secure the repository of any stored data against unauthorised users.

Another ethical point to consider is the code that is used to create this application. I should take due care to create this app using only my own code, except where express permission has

¹⁸<http://phonegap.com/>

¹⁹Phonegap as an example, uses HTML, CSS and Javascript to create an application for both platforms simultaneously

²⁰Whilst StackOverflow is not preferable as a citable source, finding an admission from Apple on this point is very difficult

²¹Unfortunately, later modifications to the application to circumvent issues with using signal strength meant this was no longer an issue

been granted. In the same vein, I must make sure to adhere to all licenses for any libraries which I use in development.

Chapter 3

Requirements

The client end of this project consists of two parts - an Android application that can determine user location (LOST) and provide route-planning inside the building, whilst feeding data to the backend, which will store the data and provide functions to view it as information. Alongside this, I will develop a secondary Android application (FOUND) so that I can easily collect wireless data to test the application with. FOUND will not be available for users of the resulting software package for this project, however when considering possible extensions after the BSc report, FOUND would be an active component to allow for the easy mapping of a building in a bespoke software package.

3.1 LOST - Android Application

3.1.1 Functional Requirements

- F1** - The application will be able to retrieve the most recent SQLite database containing wireless access point location information upon startup.
- F2** - The application will be able to accurately discern the user's location, using data provided by an SQLite database.
- F3** - The application will be able to provide the user with route-planning data, allowing the user to select a start and an end point, and provide the shortest path between the two points.
- F4** - The application will be able to discern whether the device is currently in the correct building for usage.

F5 - The application will send anonymous information on the user's location to the backend periodically, if it is in the correct building

F5.1 - The user will be able to turn off this functionality

F6 - The application will periodically check if it has a functioning internet connection - if it is in the correct building. If the internet connection is malfunctioning, it will save this data.

F6.1 - Once the application regains internet access, it will send location data, and the MAC address of malfunctioning router to the backend.

F6.2 - The user will be able to turn off this functionality.

3.1.2 Non-functional Requirements

NF1 - Application will not cause an 'ANR' (app not responding) error on any 'low-end' smartphone (at time of writing, low end specifications are considered to be 1Ghz dual core processor and 768mb RAM¹).

NF2 - Application will provide location accurate to within 20 feet.

NF3 - Application will use as little battery life as possible, using less than 5% battery life within an hour.

NF4 - Application will minimise data usage, making sure not to post data back using mobile data.

NF5 - Application will not send data that can be used to identify the user, or any information that may be personal to the user / a breach of Data Protection laws.

3.2 Backend

3.2.1 Functional Requirements

F1 - Backend will operate a RESTful API that will allow the mobile application to post data to the server

F2 - Backend will store POSTed data to an SQL database

F3 - Backend will be retrieve data from SQL database, and display to user as information

¹Figures created by comparing phone specifications at http://www.phonearena.com/news/Affordable-not-cheap-15-best-low-cost-smartphones_id58696

F4 - Backend will require user authentication to allow access to displayed data

F5 - Backend will be able to store wifi location database, and serve it to a mobile device

3.2.2 Non-functional Requirements

NF1 - Backend will provide an easily navigable user interface

NF2 - Backend will show data in an easily readable manner

3.3 FOUND - Android Application

3.3.1 Functional Requirements

F1 - Application will scan local wifi access points

F2 - Application will insert the top access point into an SQLite database as a node, along with an edge connecting it to another node in the directed graph

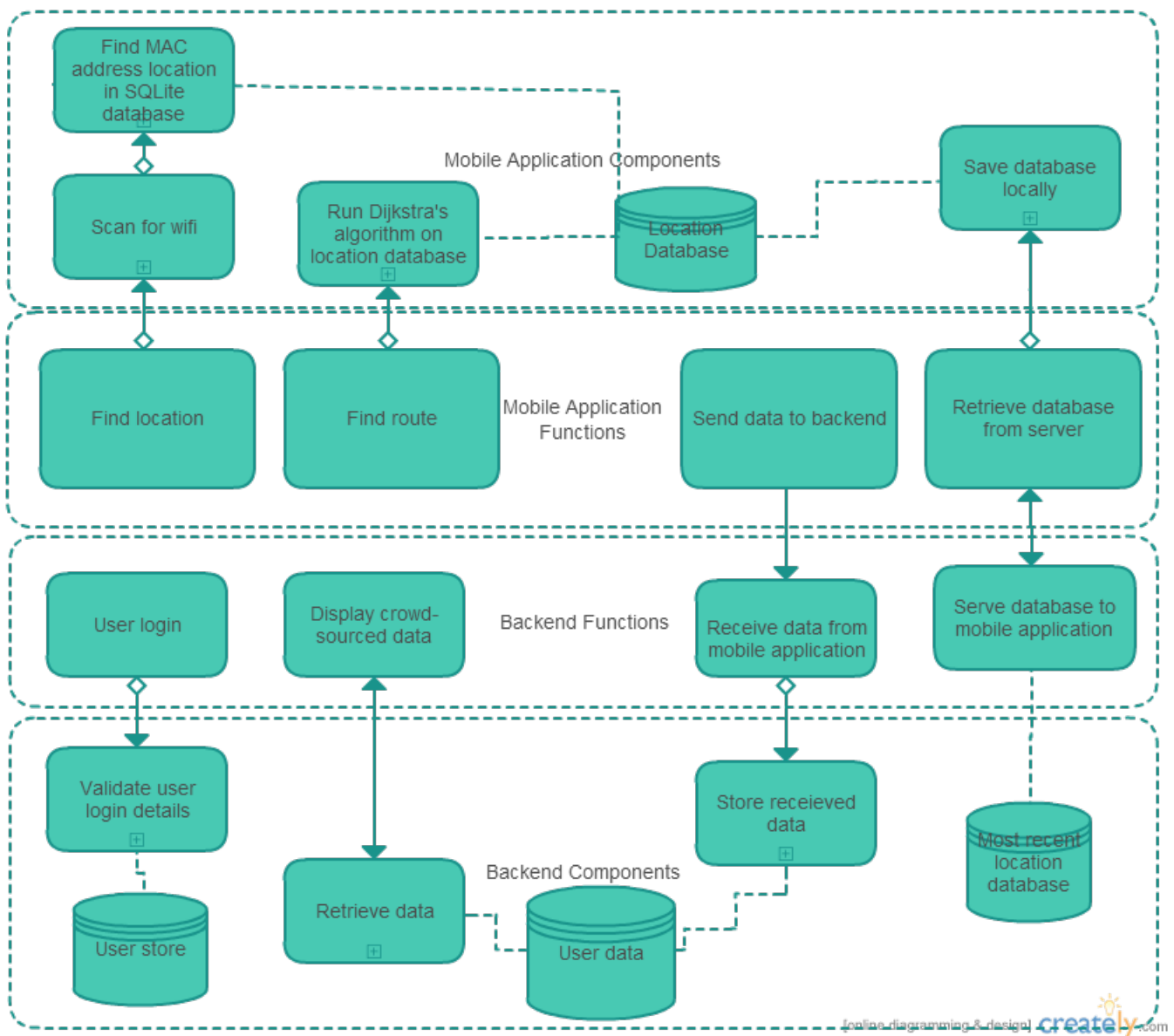
Chapter 4

Design

The application is split into two parts - therefore it will be detailed as a whole product, before the design of each component is analysed individually. Generally, the backend and the frontend application work as their own components, with the only interaction between the two being fetching the database, and sending data from the application to the database; because of this, an agile development cycle will be used.

Further to this, a basic version of the Android application will be developed first, and an initial deployment of the backend, with the ability to receive data from the mobile application after that - from then, incremental additions will be made to either component, building for functionality first, then returning later for optimisation.

4.1 Component Diagram



4.2 LOST

The 'front end' of the project, the LOST application provides most of the functionality detailed in the earlier chapters of this report. In this section, I will go into more detail as to what each function will need to do in order to provide the outlined functionality. You will find a Use Case Diagram, along with wireframes for each aspect in the Appendix.

4.2.1 Finding Location

One of the most critical aspects of the application, and also one of the most challenging as a great deal of consideration needs to be paid towards efficiency, accuracy, and precision - it will generally be more expensive in terms of resources, and in fact battery power, to increase the precision.

In order to ascertain location, the application will need to perform a scan of nearby wireless access points. The MAC address and signal strength of each access point will be collected, and put in descending¹ order of strength. The original intention² at this point was to use actual signal strength figures to get the most precise location information, as done by Cook *et al*[5].

To get this location information, the top three MAC addresses in the list would be used to filter the location database down to a single node, representing that location³. This location would then be displayed to the user, along with an option to refresh this (performing another scan, incase the user doesn't think they are in the right location). There will also be a button that allows the user to use their given location as the start or end point for route planning.

4.2.2 Route Planning

This function ties in with the location finding function of the application - the user will either be able to load it normally, or load it from the 'Location Finder' part of the application - if this is the case then either the 'to' or 'from' sections of the application will contain the user's location.

¹Or considering how signal strengths are measure, ascending. Signal strength is measured in DBm and measure as negative figures. The 'quieter' a signal is, the weaker it is - therefore a signal strength of -10 is much more powerful than a signal strength of -80

²Results of this will be detailed in the 'Implementation' section of this report.

³Discussed further in the Databases section

References

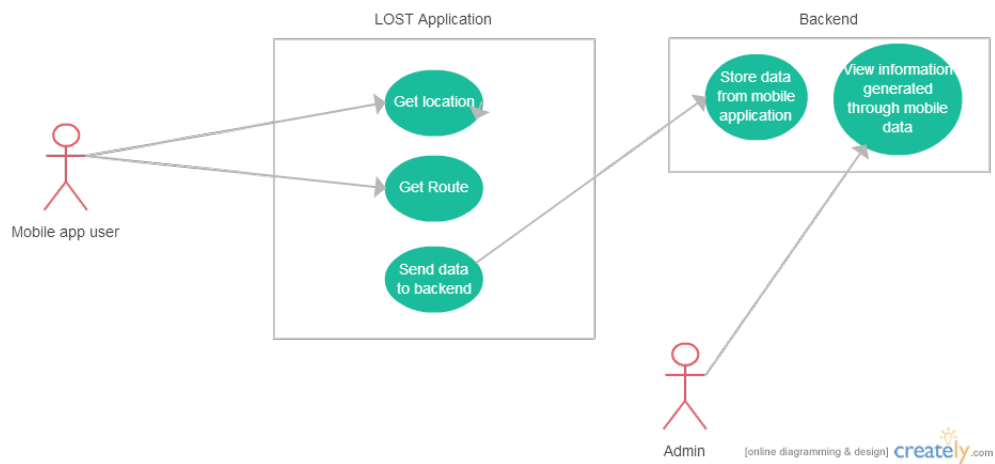
- [1] Stack overflow - iphone signal strength. <http://stackoverflow.com/questions/2959567/iphone-signal-strength>.
- [2] Stack overflow - measuring signal strength. <http://stackoverflow.com/questions/21715690/measuring-signal-strength-from-wifi-to-iphone-ipad>.
- [3] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *Handheld and ubiquitous computing*, pages 304–307. Springer Berlin Heidelberg, 1999.
- [4] Yongguang Chen and Hisashi Kobayashi. Signal strength based indoor geolocation. In *Communications, 2002. ICC 2002. IEEE International Conference on*, volume 1, pages 436–439, 2002.
- [5] B Cook, G Buckberry, I Scowcroft, J Mitchell, and T Allen. Indoor location using trilateration characteristics. In *Proc. London Communications Symposium*, pages 147–150, 2005.
- [6] Google. Configure access points with google location service. <https://support.google.com/nexus/answer/1725632?hl=en>.
- [7] IDC. Smartphone os market share, q4 2014. <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>, 2014.
- [8] King’s College London. King’s college london headcount summary 2012-2013. <http://www.kcl.ac.uk/aboutkings/facts/ida/stats/HeadcountSummary2012.pdf>, 2012.
- [9] Jakob Nielsen. The 90-9-1 rule for participation inequality in social media and online communities. <http://www.nngroup.com/articles/participation-inequality/>, 2006.

- [10] National Research Council (U.S.). Committee on the Future of the Global Positioning System, National Academy of Public Administration, Commission on Engineering, and Technical Systems. *The Global Positioning System: A Shared National Asset*. Romantic Reassessment; 118. National Academies Press, 1995.
- [11] Felix Richter. 15 years of google. <http://www.statista.com/chart/1502/15-years-of-google/>, 2013.
- [12] Dario Salvi, Manuel Ottaviano, Ignacio Peinado, and Maria Teresa Arredondo. An architecture for data collection and processing in context-aware applications. In *Adjunct Proceedings of the 3rd European Conference on Ambient Intelligence, AmI09, Salzburg, Austria*.
- [13] J. Schiller and A. Voisard. *Location-Based Services*. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 2004.
- [14] Steven J. Vaughan-Nichols. How google-and everyone else-gets wi-fi location data. <http://www.zdnet.com/article/how-google-and-everyone-else-gets-wi-fi-location-data/>, 2011.

Appendix A

Extra Information

A.1 Use Case Diagram



A.2 Wireframes