

UV Data Analysis and Navigation

Aditya Modi, Jerrid Matthews, Mario Gerla
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA 90024
{adityam, matth122, gerla}@cs.ucla.edu

1 Abstract

Pedestrians exposure to UV Radiation depends on many factors, most important of which are geographic location and environmental properties. UV radiation in moderation is beneficial to human health like production of Vitamin D. However, overexposure to UV Radiation can lead to many health problems including skin cancer. The purpose of this project is to provide user with a path from the source to the destination, that has minimum UV exposure, thus reducing the health risks. UV data is gathered using a circuit which contains sensors for UVA and UVB. Further, statistical analysis is performed on the data. Data analysis is then done to determine what the minimum number of UV samples is to assert the UV exposure on a path. The aim here is to obtain an accuracy of above 75% for UVA and above 95% for UVB. A method is also devised to estimate UV data from the neighboring road segments if the number of samples for a given road segment is lesser than the minimum number of UV samples required for analysis. A Web Service is built to cater the developers to use the system in order to get the best path to route a pedestrian with minimum UV exposure. A web application is also built to consume the web service.

2 Introduction

Scientists have come to a similarly dichotomous recognition that exposure to the ultraviolet radiation (UVR) in sunlight has both beneficial and deleterious effects on human health. UV exposure is beneficial to humans in an optimum amount since it is necessary for the production of Vitamin D in the body. Unlike other essential vitamins, which must be obtained from food, vitamin D can be synthesized in the skin through a photosynthetic reaction triggered by exposure to UV radiation [1]. So, it is necessary for humans to get an optimum amount of UV exposure. However, over exposure

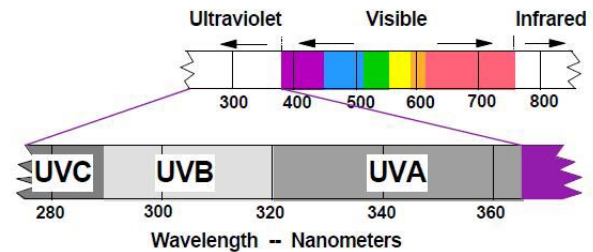


Figure 1. UV Spectrum

of UV have a lot of bad implications on the body like skin cancer and sun burns. These health effects could have been worsened due to the depletion of the ozone layer resulting in increase in UV radiation [2]. The effect on human health of solar UV at the earth's surface depends on many factors like geographic location and environmental properties.

The UV spectrum, shown in Figure 1 is divided into three wavelengths: UVA 315-400nm, UVB 280-315nm and UVC 200-280nm [4]. The ozone layer of earth blocks UVC radiations whereas UVA and UVB radiations pass through the atmosphere. UVA radiation penetrates deeply into the skin, where it can contribute to skin cancer indirectly via generation of DNA-damaging molecules such as hydroxyl and oxygen radicals. Sunburn is caused by too much UVB radiation; this form also leads to direct DNA damage and promotes various skin cancers. Erythema is caused primarily by the UVB [3]. It causes reddening of the skin by damaging the epidermal layer. The amount of UV radiation that is absorbed or scattered is determined by a number of factors. For measurement of UV Radiation, there is a metric called UV index [5].

Since over exposure of UV on humans has several adverse effect, exposure to UV radiation should be avoided. The paper highlights a web service to route users through routes with minimum UV exposure and thus reducing the amount of UV exposure on pedestrians. It also showcases a prototype to consume the web service by building a web application.

In order to estimate the UV levels on a route, analysis was done in order to determine minimum number of samples necessary to assert the UV levels on a particular route. The analysis plays a crucial role in the navigation application since there needs to be a certain confidence in estimating the

UV levels on a particular route. Another interesting problem that is discussed in the paper is to estimate the UV on a route when the number of samples are less than the actual number of samples required. The UV level in these cases is estimated from the UV levels of the neighbors. Using the results mentioned in [6], it is reasonable to estimate the UVA accuracy to 75% and UVB accuracy to 95%.

3 Approach

Work is mainly divided into the following parts:

- Route Selection
- Sensor Device Selection
- Data Collection
- Data Analysis
- Web Service Development
- Navigation Web Application

For the project to be a success, the most crucial part is selection of the route for the experiment. The route that is selected needs to have a good mix of trees, buildings and open spaces. A major criteria was to also select a path which has alternate routes so that decision on suggesting a route can be taken by comparing the options based on the UV levels. After the route selection, a selection was to be made for selecting the devices to collect the UV data. A requirement for sensor selection was that device should have sensors to measure both UVA and UVB data. Also there needs to be an interface to transfer the data collected to the computer. A GPS device is also necessary to keep a track of the latitude and longitude points where the UV data is collected.

After the data was collected, analysis was done on the data to determine the minimum number of samples necessary to assert the level of UV on a particular route. After the analysis is completed, a web service is developed to incorporate the analysis done so far, and help the developers to create platform independent navigation system on top of the system created. The main idea being the users of the application uploading data to the common server and the web service provides an API to the external world developers. Finally, a web application was also created to consume the web service and provide as a basis for other application developers.

4 Implementation

4.1 Route Selection

Selecting a route was extremely crucial for the project since UV exposure is affected by geographic and environment properties like trees and buildings. A route should be selected with a good balance of open space, trees, buildings, and a mixture of all previously mentioned objects. After gauging different route options, a route was selected around UCLA area which had a mix of open space, buildings and trees. Our source was 606 Levering Ave and destination was 11020 Kinross Ave 2. The alternate routes were as follows

- Via Veteran Ave
- Via Levering Ave and Gayley Ave
- Via Weyburn Ave

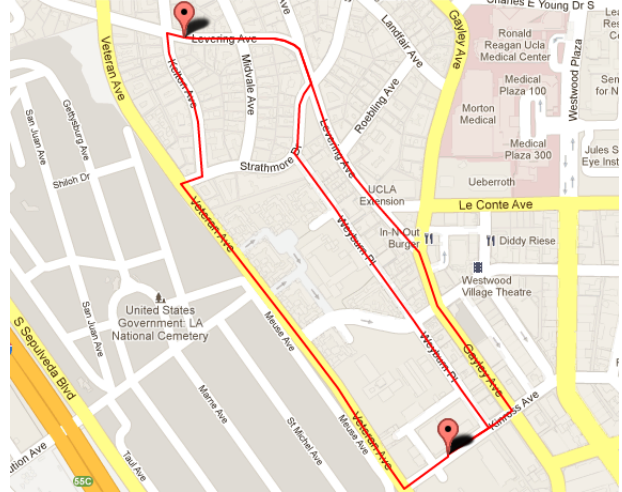


Figure 2. Routes taken for doing the analysis.

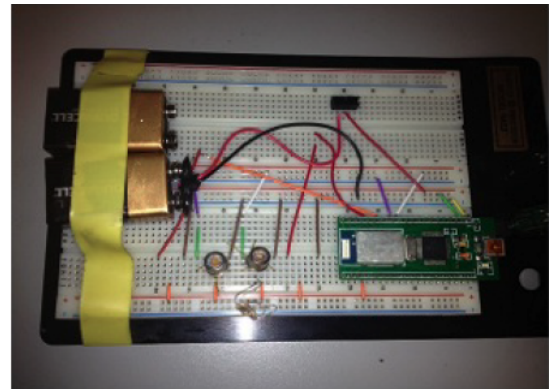


Figure 3. Sensor to collect UV Data.

Veteran Ave route had a good mix of buildings and trees, while Levering Ave/Gayley Ave and Weyburn Ave routes had a good mix of buildings and open spaces.

4.2 Sensor Device Selection

Data collection had two main components, UV readings and GPS readings. For UV Readings, a sensor built at UCLA was used 3. The reason for selecting this sensor was the fact that all our requirements were satisfied. The device had 2 sensors for measuring UVA UVB each. The units of data output is mW/nm^2 . The device also had a bluetooth which facilitated transfer of data from the device to a computer. Since there was no GPS device built on the UV sensor, Android device was used to collect GPS readings. After gauging various options for a GPS software on the android market, GPS logger software 4 was finally selected to log the GPS data since it was very flexible and provided different options to log the data. GPS logger used to log the GPS readings, which were at a later stage merged with the UV sensor readings to get the UV exposure at specific latitude and longitude points.

4.3 Data Collection

For collecting the readings, actual walks were done on the selected route 2. The walks were done by holding the sen-

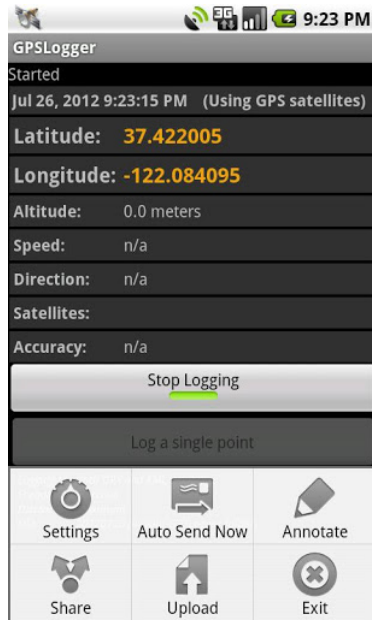


Figure 4. Software to collect GPS Data.

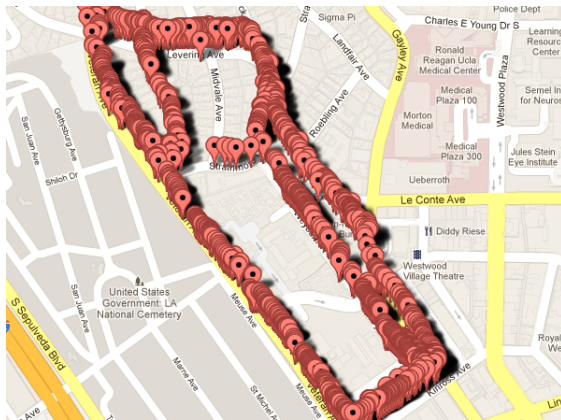


Figure 5. Data points with UV Readings.

sor device, a laptop to transfer the readings and an android phone to collect the GPS readings. Readings were taken on October 16th 2012 between 9:05a.m. and 9:42a.m., October 23rd 2012 between 9:14a.m. and 9:46a.m., November 3rd 2012 between 9:11a.m. and 10:05a.m., November 10th 2012 between 9:24a.m and 12:12p.m. with clear sky. Protection from the sun rays may be different on different sides of the road depending on height of the buildings, presence of trees, time of days etc. Thus, readings were taken by walking on both sides of the road wherever possible. Figure 5 shows a plot of points where readings were taken by walking along the road.

4.4 Data Analysis

4.4.1 Data Cleansing

Data cleansing is the most important step of any project where data analysis is involved since analysis can only be performed on clean data and that noisy data should be removed. There were UV readings taken by one device and the

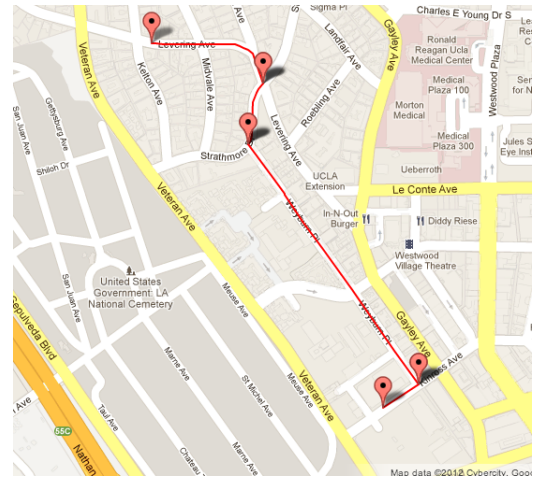


Figure 6. Steps in a single Leg.

GPS readings were taken by another device. So, a script was written to merge the readings so that exact UV readings were collected on the logged GPS points. Also, there were couple of sensors for both UVA and UVB each. So the data was analyzed to determine which of the two sensors readings were more reliable. After careful analysis, one of the two sensors was taken into consideration. There were also some random erroneous readings reported for the sensors sometimes which were discarded. After performing all of the above mentioned steps, the data was clean and ready for analysis.

4.4.2 Determining minimum number of sample points

The main aim of the project is to suggest the best route to the user out of the many route options he could take. After carefully considering many options online, Open Street Maps [9] and Google Maps [7] were the good options to be considered. The choice of selecting Google Maps was the fact that no efforts would have to be made for calculating the actual navigation route since there are many robust API's available which good be directly used out of the box. The only programming effort was required to parse the data and selecting only the information which was relevant for the project.

In order to determine the various alternate route options that can be taken to travel from source to destination, Google Maps [7] data is used. Open Street Maps option was also considered. Google Maps returned alternate route options when source and destination locations were passed where as in Open Street Maps, the output was in form of the nodes en route to the destination. There is a database of nodes on any path which is returned by Open Street Maps. Since alternate paths were necessary in this project, google maps was used. Google Maps returns the various routes that can be taken which are called as *Legs* and the various sections of the roads that changes are called as *Steps*. In the route selected for the experiment, each of the red lines in Figure 2 indicate each leg and each of the red lines between the markers in Figure 6 indicate each step. Minimum number of points needed to assert the UV level with a certain confidence was to be found. Analysis was initially done by taking one step and finding the number of points required on that step. The steps

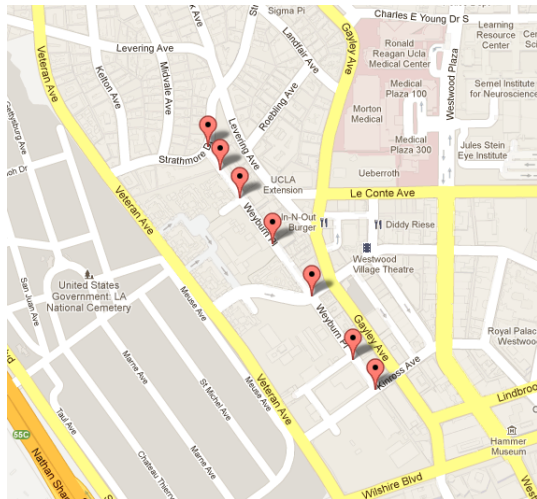


Figure 7. *Segments in a single Step.*

that Google Maps returns are too large to do efficient analysis, therefore each step returned by Google Maps needs to be broken down into smaller segments. The steps returned by Google Maps had distances which were not consistent. Also, if we consider large steps, then there may be data at only one end of the step which may not be co-related with data at the other end of the step. Also, there may be trees/buildings on one end and not on the other. So, if steps are broken down, it increases the likelihood of having more accurate estimates of the pedestrian UV exposure along the step. After analyzing various options the Open Street Map API [10] was used because it breaks the steps down into smaller pieces. On an average, the segment size was around 200 to 300 ft. When passed two end points, the API returns all the nodes in its database which come in between the end points. Thus, a pair of nodes would form a segment. Again, scripts were written to parse the data and extracting the information required for the project. Figure 7 indicates segments within a single step.

Data is analyzed to determine the minimum number of points required to get the accuracy of UVA above 75% and accuracy of UVB above 95%. The accuracy limits are selected from [6] since extensive research was done and also the sensor used was the same one as used in this experiment. Bootstrapping method is used to determine the minimum number of points. Bootstrapping works as follows:

1. Take the average of all the points on a particular segment which would become our actual average
2. Select one point randomly
3. Calculate the accuracy with respect to the actual average
4. Select one more point randomly without replacement since the result should not be biased and accidentally choose the same data points
5. Take the average of all the points selected till now
6. Calculate the accuracy with respect to the actual average
7. Repeat from Step 4 till desired accuracy is achieved

Algorithm for Bootstrapping

Listing 1. R Code for Bootstrapping

```
route <- read.csv("route.txt", header=
  TRUE , sep=",")
for(noOfSample in 1:10){
  trueMean = mean(route$uvb)
  uvb = route$uvb #getUVB Readings
  uvbReadingsVector <- vector()
  for(i in 1:10){
    uvbSample = sample(uvb,noOfSamples,
      replace=FALSE)
    sampleMean = mean(uvbSample)
    accuracy = (100-abs((trueMean-
      sampleMean))/ sampleMean*100)
  }
}
```

Bootstrapping method is chosen since the distribution is completely random and the readings along the segments are not consistent. Bootstrapping is the best method in these cases. For analysis variety of segments are covered by considering different road segments at different orientations and different length of segments.

Detailed analysis of one of the segments that was considered is as follows. A segment from Veteran Avenue was selected 8. The selected segment had a trees on one side of the road and buildings on other side of the road. Start Point was 34.060067, -118.449498 and End Point was 34.058945, -118.44864. Total length of the segment was 480 ft.

No of Readings	UVA Accuracy	UVB Accuracy
1	72.03039	93.37286
2	72.84915	93.7059
3	73.27921	94.36731
4	78.09173	94.94913
5	76.75012	95.06996
6	77.60953	97.47088
7	78.30137	97.91901

Table 1. Accuracy table for bootstrapping

As observed from the table that UVA and UVB take around 4 readings to get the desired accuracy. After repeating this experiment for many different segments, observation was made that in the worst case, 6 points on a segment are necessary to get accuracy above 75% for UVA and 4 points on a segment are necessary to get accuracy above 95% for UVB. In average case, 4 points on a segment are necessary to get accuracy above 75% for UVA and 3 points on a segment are necessary to get accuracy above 95% for UVB.

4.4.3 *Devising a method to handle segments with less number of samples than minimum required*

There may be some segments, which may not have the minimum number of readings that is necessary to assert the UV values like the segment in the Figure 11.

The problem of some segments not having desired number of readings is countered in the way mentioned below. In

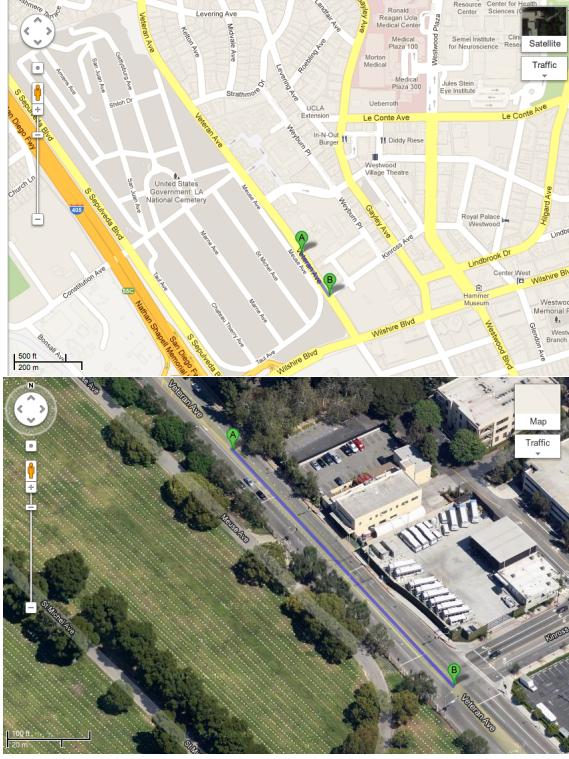


Figure 8. Sample selected segment map view

that case, a weighted average of the readings from the neighboring segments is taken. The weight is in terms of how far the segment is from the segment, which is currently under consideration and also, the number of readings that the current segment has i.e. the segment which is closer to the segment under consideration and has more readings is given more weight than the segment otherwise since there would be more confidence in the closer segments than segments which are not close relatively. The immediately neighboring segments have a weight of 0.5, the weight is decreased by 0.1 as segments farther from the segment under consideration is taken into account for calculating the UV values. This method was selected since segments, which are near the segment under consideration, have a high correlation of readings with the segment. The window of neighboring segments is expanded till the desired number of readings for the segment is achieved. For instance the green segment in Fig 10 is taken into consideration. Since it does not have any UV readings, the UV Values need to be taken from the neighboring segments. As per the technique discussed, it will try to find readings from the immediate neighbors, since they are not enough, it will try to expand further till it reaches the segments with 4 readings, when the total number of readings would be greater than the minimum readings necessary. The UV exposure in the green segment would be calculated in Equation 1

$$UV = \frac{0.5 \times 3 \times 300 + 0.3 \times 4 \times 200}{0.5 \times 3 + 0.3 \times 4} \quad (1)$$

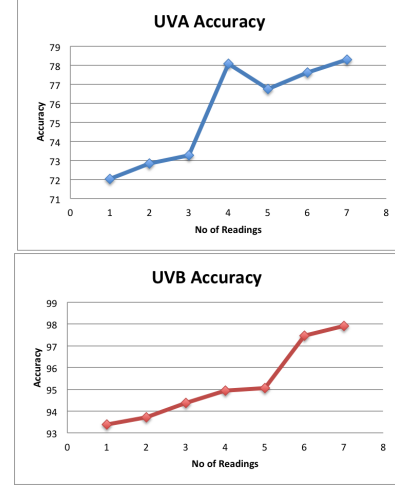


Figure 9. Accuracy v/s No of Readings

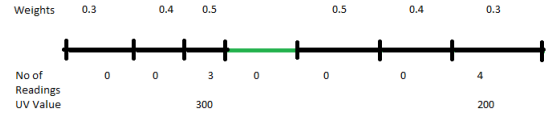
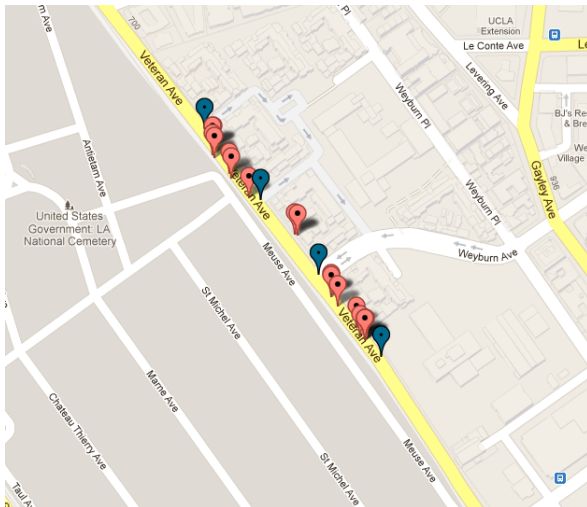


Figure 10. A sample segment with less number of readings

For instance, if the segments in Figure 11 is considered. The segments end points are marked by blue colored markers and the red color markers indicate readings at those points. As observed in the figure, the number of readings are less than the minimum number of readings required. So, a weighted average of the neighboring segments is taken. Since both the segments are immediate neighbors, weight would eventually be proportional to the number of readings in each segment as there would be more confidence in considering readings from the segment that has more readings as compared to considering the segments with less number of readings.

4.5 Web Service Development

Web Service was developed to provide API to the developers to make use of the common platform and develop navigation applications based in UV data. The Web Service developed is similar to the Google Maps Web Service [8] with extra information about the UV levels along the route. Initially, data is pulled from Google Maps where information about routes, distance between the source and destination, time to travel and most importantly alternate route information is pulled. After parsing the information, we get the alternate routes i.e. the *legs*. For each *leg*, information about *steps* in each route is also retrieved from Google Maps API. For each *step*, Open Street Maps API is called to get the segments inside each step. Then, UV data about each segment is retrieved from the database. After calculating the UV levels on each of the paths, an output of a path with minimum UV levels is passed to the caller. Thus, data from Google Maps is joined with data from Open Street Maps and joined with the UV data on a path and returned to the caller. A choice was to be made whether to display the output in XML format or



JSON. Finally, the output was in JSON format since its easier to read. Its also less decoding effort to programmers as compared to XML output. The inputs to the web service are source latitude, source longitude, destination latitude, destination longitude and UV criteria. UV criteria is whether route with minimum UVA or minimum UVB or an average of UVA and UVB is required. A part of sample output from the webservice is in Fig 12

4.6 Navigation Web Application

A Web Application is built in order to consume the web service built. The purpose of the web application is to serve as a basis for developers to use the system and build applications on top of it. The web application shown in Fig 13 is built using the web service and google maps javascript API to render the maps on the web page. Similarly other desktop and mobile applications can also be built. When entered the source and destination latitude and longitude and selecting the UV criteria, the application will suggest the most optimum route based upon the selected criteria as shown in Fig 13.

5 Experiment

Taking the results from the data analysis that was done, experiments were performed to verify whether the desired accuracy was achieved. As per the observations from the data analysis, in the worst case, 6 points on a segment are necessary to get accuracy above 75% for UVA and 4 points on a segment are necessary to get accuracy above 95% for UVB. In average case, 4 points on a segment are necessary to get accuracy above 75% for UVA and 3 points on a segment are necessary to get accuracy above 95% for UVB. The best possible approach to test the analysis was to randomly take the number of sample points actually necessary and check whether the desired accuracy is achieved. To remove the randomness from the process, sample number of points were taken around 100 times and their accuracy was calculated. This accuracy was verified with the desired accuracy.

Again, some segments were selected. Accuracy was determined by considering the worst case as well as average

case scenario. For instance, a segment in Figure 8 was considered. The accuracy in the worst case where 6 samples for UVA were taken and 4 samples for UVB was taken gave an average accuracy of UVA at 79.51% and UVB at 96.75%. When average case where 4 samples for UVA were taken and 3 samples for UVB was taken gave an average accuracy of UVA at 75.93% and UVB at 96.51%. From the observation of all the readings, it could be inferred that if 6 samples to estimate UVA and 4 to estimate UVB are taken, then the accuracy remains above the thresholds that are required.

Listing 2. R Code for verifying accuracy

```
route <- read.csv("route.txt", header=
  TRUE , sep=",")
sampleSet = 4 # for worst case scenario
uvb = route$uvb # get UVB values
uvbAccuracy <- vector()
for(i in 1:100){
  uvbSample = sample(uvb, sampleSet ,
    replace=FALSE)
  trueMean = mean(route$uvb)
  sampleMean = mean(uvbSample)
  t = abs(100-abs((trueMean-sampleMean)/
    sampleMean*100))
  uvbAccuracy[i]=t
}
```

6 Future Work

The work can be extended to do analysis on different paths. That way, the current number of samples of can be reaffirmed. Since the accuracy of UVA is aimed to be 75%, a more accurate UVA sensor could be used and a higher UVA accuracy could be aimed for. The current accuracy limits are selected from [6] since extensive research was done and also the sensor used was the same one as used in this experiment. UV sensors could be interfaced directly with the android device which could facilitate more flexibility to gather data. After that is achieved, more users could be encouraged to gather UV data. In this way, crowd sourcing could be used to populate the database eventually helping users navigate more routes.

7 Conclusion

It could be concluded that in the worst case, 6 points on a segment are necessary to get accuracy above 75% for UVA and 4 points on a segment are necessary to get accuracy above 95% for UVB. In average case, 4 points on a segment are necessary to get accuracy above 75% for UVA and 3 points on a segment are necessary to get accuracy above 95% for UVB. The number of points on a segment is not dependent on the length of the segment since the length of the segments are generally small. A web service was developed to estimate UV levels on a path and route the user through the path having the minimum UV value.

8 References

- [1] M. Nathaniel Mead, Benefits of Sunlight: A Bright Spot for Human Health
- [2] Richard H. Grant and Gordon M. Heisler, Effect of Cloud Cover on UVB Exposure Under Tree Canopies: Will Climate Change Affect UVB Exposure?
- [3] Richard H. Grant , Gordon M. Heisler and Wei Gao, Estimation of Pedestrian Level UV Exposure Under Trees
- [4] <http://articles.mercola.com/sites/articles/archive/2012/03/26/maximizing-vitamin-d-exposure.aspx>
- [5] Wong, C.F., Toomey, S., Fleming, R.A., Thomas, B.W., 1995, UV-B radiometry and dosimetry for solar measurements, Journal of the Health Physics Society, Vol. 68.
- [6] Jerrod Matthews, Farnoosh Javadi, Gauresh Rane, Jason Zheng, Giovanni Pau, Mario Gerla, Ultraviolet Guardian - Real Time Ultraviolet Monitoring
- [7] <http://maps.google.com>
- [8] <https://developers.google.com/maps/documentation/directions/>
- [9] <http://www.openstreetmap.org/>
- [10] <http://wiki.openstreetmap.org/wiki/YOURS>