Cost-Effective Positioning based on WiFi-Probes: A Quantitative Study

Deriving the Position of a Smartphone using the Signal Strength of WiFi-Probes

Alexander Ljung
Hannes Knutsson

Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden
This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science. The thesis is equivalent to 20 weeks of half time studies.

The authors declare that they are the sole authors of this thesis and that they have not used any sources other than those listed in the bibliography and identified as references. They further declare that they have not submitted this thesis at any other institution to obtain a degree.

Contact Information:
Authors:
Alexander Ljung
E-mail: alexander@lju.ng

Hannes Knutsson
E-mail: hannes.knutsson96@gmail.com

University advisor:
Dr. Prashant Goswami
Department of Creative Technologies

Faculty of Computing
Blekinge Institute of Technology
SE–371 79 Karlskrona, Sweden

Internet : https://www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57
Abstract

In the modern society, almost everyone has a smartphone. These devices tend to almost always use WiFi-networking. For the device to identify nearby WiFi access points it has to send out WiFi probing broadcasts. Nearby access points respond to these broadcasts in order to let the device know that they are within reach. This technique is called active scanning. This paper aims to answer if it is possible to use the signal strength of these broadcasts to localize the device transmitting them. We are interested in the possibility of creating this kind of system and the accuracy that it would be able to provide.

This is a quantitative study where we produce our results based on experiments, measurements and observations. The experiments are set in a large square shaped area. A sensor was placed at each corner of the area that the smartphone will be tracked within. The smartphone will be sending WiFi probing broadcasts that will be monitored and measured by the sensors. The strength of the broadcast signal will be converted into the relative distance between the devices position and the sensors. These four distances, collected from each of the sensors, will further be converted into a position within the area by using trilateration. To measure the accuracy of the system, the true position of the device will be compared against the calculated position from the system using only the signal strength. Further, a deviation in the distance between the two locations will be calculated.

The experiments resulted in a positioning system that was able to estimate positions within an 80 x 80m area. Fourteen location positions were taken which resulted in a mean deviation of 16.6 meters from the true location and a root mean squared error of 19.5 meters. We concluded that more readings within the same position gave a significant increase in accuracy, to the expense of time. Using single measurements would be more practical, but would not produce reliable positions.

Keywords:
WiFi, Probe Broadcast, Local Positioning System, Trilateration, RSSI.
Acknowledgments

We would like to thank Dr. Prashant Goswami for leading us throughout the process of conducting these experiments and the writing of this paper. We would also like to thank Tomas Petersson for helping us understand the mathematics involved in this project and reasoning with us regarding how to proceed with the tests.

Thank you both for the help you so kindly provided us with.
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Visual presentation of a single sensor (S) interpreting the signal strength received from the probing broadcast of the smartphone to be tracked (P). A stronger signal from the smartphone would result in a shorter radius (d) circle, whilst a weaker signal would indicate a larger distance, thus resulting in a larger radius circle. The circle represents the possible locations that the smartphone could be located on. The sensor can not individually tell exactly where on the circle’s edge the smartphone is located.</td>
<td>6</td>
</tr>
<tr>
<td>3.2</td>
<td>Representation of the testing area (<em>large black square</em>), the smartphone to be tracked (P), the sensors (S&lt;sub&gt;1&lt;/sub&gt;, S&lt;sub&gt;2&lt;/sub&gt;, S&lt;sub&gt;3&lt;/sub&gt;, S&lt;sub&gt;4&lt;/sub&gt;) and their individual understanding of where the smartphone is located (<em>circles around sensors</em>).</td>
<td>7</td>
</tr>
<tr>
<td>3.3</td>
<td>Panorama photograph over the area used for testing. One of the sensors can be seen on its tripod in the middle of the image close to the camera. In the middle of the area, Alexander Ljung can be seen operating the server in preparation for a test.</td>
<td>8</td>
</tr>
<tr>
<td>3.4</td>
<td>The four combinations of sensors possible for using the trilateration technique. These positions can be combined into one more accurate position.</td>
<td>10</td>
</tr>
<tr>
<td>3.5</td>
<td>Visual presentation of positions within the testing area that were used for testing the accuracy of the system.</td>
<td>11</td>
</tr>
<tr>
<td>4.1</td>
<td>Decay of signal strength when the distance from the source is incremented. The yellow line represents the strongest signal strength recorded, the red line represents the weakest signal strength recorded and the blue line represents the average signal strength recorded. 10 measurements were recorded at each ten-meter distance from the sensor. The leftmost data point is recorded at the distance of one meter from the sensor.</td>
<td>13</td>
</tr>
<tr>
<td>4.2</td>
<td>The result of putting the RSSI-values (previously displayed in figure 4.1) into the function 3.1. The red line represents the highest distance calculated for a single data point, the yellow line represents the shortest distance calculated for a single data point and the blue line represents the calculated distance based on the mean signal strength. The green line is added as a reference to what perfect conversion would look like.</td>
<td>14</td>
</tr>
</tbody>
</table>
4.3 Deviation from the true distance of the calculated distances from figure 4.2. The red line represents the highest distance calculated for one single data point, the yellow line represents the lowest distance calculated for one single data point and the blue line represents the distance calculated from the mean signal strength.

4.4 The fourteen tests of thirteen different locations within the area. The red dot represents the position the smartphone was located at. The Blue dot represents the position calculated by the system based on the signal strength of probing broadcasts.
Chapter 1

Introduction

Our society has become extremely dependent on technology. This shows in almost any workplace, home or school. Although the list of where you may find technology can be made very long, the point to be made is that technology is everywhere. In the last 10 - 20 years it has become way more common for people to have personal cellphones or smartphones as they are called. The name smartphone is used since the phones nowadays are so much more than just telephones. Smartphones can do almost anything, it is a powerful pocket-sized personal computer. These devices have to communicate with the internet to be able to utilize their full potential. This is usually achieved in a couple of different ways.

One of the most common ways for a smartphone to access the internet is via WiFi[14]. Almost anyone you meet is aware of WiFi technology and how to use it in their home or work environment. It is generally not hard to set up and get working, but when it comes to how it really works; how many of these people do you think understand the actual underlying technology?

A smartphone with WiFi activated (sometimes not necessarily even activated[13]) constantly looks for access points to connect to. To be fair, WiFi would be useless if there were no access points for clients to connect to and access the internet. To check whether there are any nearby access points, the smartphone broadcasts a wireless probing message out in the open. This can occur as frequently as 2000 times per hour, depending on the ongoing activities within the device[9]. The smartphone at this stage has no idea at all whether there are any nearby access points to connect to or not. This probing message could be looked upon as a cry in the dark. Access points that receive this probe will answer the smartphone with details about its presence. This way the smartphone will able to ask for and find new access points for its user to connect to. This is a process called active scanning[18][9].

What if someone else, other than an access point were to catch these probes? What would they be able to see and what could they do with that information? The answer is both straightforward and complicated. The content of probes is usually the MAC-address of the smartphone and the name of the networks the smartphone previously has been connected to[12]. This information might not seem like much, but this has been proven to be very useful within many different areas[15][12].

When sniffing for WiFi-probes, there exist many different data metrics that can be collected. What has interested us in this study is the combination of uniquely
identifiable MAC-addresses and the received signal strength (RSS or RSSI). These two metrics combined could enable potential sniffers to estimate the distance to the smartphone by applying mathematics to the RSSI value. Individual sensors are unable to position the smartphone by themselves. They can only calculate the distance to the smartphone. Multiple sensors combined though would together be able to decide a location where the smartphone should be located, based on their individual calculated distances. For this to work, the sensors have to be able to record the signal strength of these broadcasts and convert them into distances in an accurate enough way. Depending on how accurate these distances are, the actual calculated position becomes more or less accurate when using trilateration.

Similar studies have been done multiple times before[18][16][4]. Our take on this is whether it is possible to create a positioning system for WiFi probes with easily acquirable and cheap hardware. Lots of papers use very nice equipment which users might not be able to acquire because of high prices, availability or its complexity. Since the user of the smartphone being tracked does not have to install any software or change any setting for this to work, this positioning system could have the potential to become very privacy invasive.

As this system is affordable and relatively easy to set up with some basic knowledge around the area, one could debate the ethical aspect of the potential widespread use of such a system.

The question we aim to answer in this thesis is whether it is possible to use cheap and easily acquirable hardware to create a WiFi-probe positioning system, and if so, how high accuracy could this system provide?
Chapter 2

Related Work

Many studies have been done within the area of WiFi-probes and other information that WiFi-enabled devices transmit to be able to communicate with access points. Much of the work done within this area is related to calculating positions of individuals or vehicles inside offices and storage buildings\cite{7}\cite{16}. Many papers have come to the conclusion that WiFi-probe signal strengths are unreliable and unstable when measured within a inside environment\cite{5}\cite{16}\cite{11}\cite{16}.

Even though not as many papers have been written about outside positioning using WiFi-probes, a number of studies have still been made within the area\cite{18}. This is primarily due to the fact that GPS is favored in outside conditions because of its excellent coverage and accuracy\cite{8}. Still, an alternative approach to outdoor positioning would be interesting.

Julien Freudiger with his study *How Talkative is your Mobile Device?*\cite{6} sheds light upon how frequently mobile devices are broadcasting WiFi-probes. They reach the conclusion that the frequency varies greatly depending on the operating system and its version, whereas Android generally has the highest frequency. This lays the foundation for our possibility for position tracking in the regard that a faster and more reliable position can be achieved if the device broadcasts its probes more frequently. In consideration of their studies conclusions that iPhones could randomize their MAC-address and that Blackberrys might not even broadcast probe requests, leads to our proposed system being rendered non-functional under such circumstances.

Much research has been done in the area of merely identifying devices by WiFi probe broadcasts and thereby draw conclusions. For example, a paper estimating the number of passengers on a vehicle has been written by Woramate Pattanussorn et al.\cite{15}. Adriano Di Luzio et al. explored the possibilities of de-anonymization of large crowds\cite{12} and demonstrated the possibilities in the extent to the area.

A somewhat similar yet different study to ours by V. Acuna et al. examined the possibility to utilize unmanned aerial vehicles carrying WiFi sniffers in order to localize victims in search and rescue scenarios\cite{1}. It is a more mobile setup which locates persons within specific zones using machine learning techniques. Our system will be more stationary, but hopefully able to provide a more specific position.

In the study conducted by Ryan J. R. Thompson et al. they achieved a root mean square error of 10 meters when locating an unknown source outside. This was
based on the received signal strength from 5 sensors in a 150 m\(^2\) area. The system we aim to create in our study varies from the system in the study by Ryan J. R. Thompson et al. in the way that we will be using very cost-effective hardware as sensors and will be covering an area about 40 times as large.

We aim to compare the accuracy of our system to the accuracy of the system by Ryan J. R. Thompson et al. to get a comparison of how accurate our system has become.
Chapter 3

Method

In order to investigate our research question regarding the possibility of building a positioning system based purely on WiFi probes, the following theory, implementation and execution of the system had to be defined. A successful execution should be able to generate positioning data, which in turn answers our follow up research question; what accuracy is this system able to provide?

This will be an empirical study conducted in a quantitative manner, as the data is directly generated throughout the experiment and after that manipulated into a position.

3.1 Theory

3.1.1 WiFi probes

WiFi signals can be viewed in different kinds of networking layers. When viewed as radio-waves in the physical layer, it is unnecessarily difficult to identify the transmitter since there is no easily accessible MAC-address in this layer. When instead viewed on the data-link layer, it is easy to identify the transmitter by their MAC-address\[^{15}\]. The disadvantage of using the data-link layer is that you will no longer be able to get the exact signal strength in decibel meters. By using the data-link layer, you will end up with the signal strength defined in RSSI which is short for Received Signal Strength Indicator.

RSSI values are representations of signal strength. The range in which the values may vary change from device to device based on the receiving device’s wireless networking card\[^{11}\].

3.1.2 Deriving distances and positions

The radio-waves of WiFi decrease in amplitude as the distance between the transmitter and the receiver increases\[^{3}\]. This enables a potential receiver of these signals to predict an approximate distance to the transmitter based on the strength of the signal. This can be done by using the following formula\[^{17}\][19]:

\[^{5}\]
\[ d = 10^{\left(\text{ptx} - \text{rssi}\right)/(10 \times n)} \] (3.1)

\( d \) = distance (meters).
\( \text{ptx} \) = signal strength at 1 meter.
\( \text{rssi} \) = current signal strength.
\( n \) = environment (1.5 \( \rightarrow \) 5.0, obstacles raise the value)

By converting the signal strength with formula (3.1) into a distance, the receiver will only know the distance between itself and the transmitter. This can be visualized as a circle around the receiver with the radius of the derived distance.

Figure 3.1: Visual presentation of a single sensor \( (S) \) interpreting the signal strength received from the probing broadcast of the smartphone to be tracked \( (P) \). A stronger signal from the smartphone would result in a shorter radius \( (d) \) circle, whilst a weaker signal would indicate a larger distance, thus resulting in a larger radius circle. The circle represents the possible locations that the smartphone could be located on. The sensor can not individually tell exactly where on the circle’s edge the smartphone is located.

If instead three or more devices would measure the signal strength at different locations, it would be possible to derive a position by using trilateration techniques\[4\][20]. The following functions can be used to convert three calculated distances into a \((x, y)\) position of the transmitting device:

\[ x = \frac{r_1^2 - r_2^2 + d^2}{2d} \] (3.2)

\( x \) = x-position
\( r_1 \) = distance from sensor 1 to transmitting device
\( r_2 \) = distance from sensor 2 to transmitting device
\( d \) = \( r_2 \)’s offset in x-axis from the origin
3.2 System implementation

\[ y = r_1^2 - r_3^2 + i^2 + j^2 \frac{x}{2j} - \frac{i}{j} x \]  

(3.3)

\( y \) = y-position  
\( r_1 \) = distance from sensor 1 to transmitting device  
\( r_3 \) = distance from sensor 3 to transmitting device  
\( i \) = sensor 3’s offset in x-axis from the origin  
\( j \) = sensor 3’s offset in y-axis from the origin

In both functions 3.2 and 3.3 sensor 1 is considered to be the origin of the cartesian coordinate system. All units are defined in meters.

3.2 System implementation

To test this theory, a wide, flat and open area without any obstructions was required to avoid as much interference on the WiFi-signals as possible. A large gravel yard was therefore chosen to conduct the tests at. An area of 80 x 80 meters \((6400 \text{ m}^2)\) was measured and marked. This testing area served as the range of where the smartphone was going to be located within.

Figure 3.2: Representation of the testing area \((\text{large black square})\), the smartphone to be tracked \((P)\), the sensors \((S_1, S_2, S_3, S_4)\) and their individual understanding of where the smartphone is located \((\text{circles around sensors})\).
One sensor was placed in each corner of the testing area at the height of about 90 cm. These sensors were used to record the MAC-address and signal strength of the smartphone from the probe broadcasts. As mentioned in the theory, the sensors by themselves are only able to decide an approximate distance to the smartphone. When these distances are put together from every sensor, it provides quite a clear image of where the smartphone most likely is located. This is visualized in figure 3.2 as all the sensors circles intersect at the point of the smartphone.

Not shown in figure 3.2 is the server that will collect the distances from the sensors and process it into a location within the testing area.

Figure 3.3: Panorama photograph over the area used for testing. One of the sensors can be seen on its tripod in the middle of the image close to the camera. In the middle of the area, Alexander Ljung can be seen operating the server in preparation for a test.

### 3.2.1 Equipment

The following equipment was used to conduct the experiments. All of the following items can be substituted by alternative equipment serving the same purposes.

#### Sensors

As our sensors, we chose four Raspberry Pi Zero W’s running the operating system Raspbian Lite. These mini computers were chosen to serve as our sensors based on their portability and low-cost. They were set up to work autonomously, constantly reporting gathered data to the server.

The Raspberry Pi’s were using their built in WiFi/Bluetooth chips to collect the probes and their signal strengths. No external sensors or adapters were used at all. The model of the wireless chip on the Raspberry Pi’s is Broadcom BCM43438.

#### Server

A laptop running Ubuntu 17.10 was used as the collector of the four sensors data. The server ran a WiFi-access point which the sensors connected to in order to transmit their measurement data.
3.3. Execution

Smartphone

An Android smartphone of the model Samsung Galaxy A5 (2017) was used as the device to be located.

Tripods

Tripods were used to raise the sensors approximately 90 cm from the ground, as the sensors receiving capabilities were observed to be severely limited when located too close to the ground. These tripods were constructed by floral sticks taped together.

3.2.2 Airodump-ng and other sensor software

Airodump-ng from the Aircrack-ng suite of tools (link in Supplemental Information) was used on the sensors as the software to collect the WiFi-probe broadcasts. The output from airodump-ng was redirected into a script that filtered the data based on the MAC-address of our smartphone. Our script automatically submitted the data to the server after the filtering.

A patch was installed on the sensors since they originally did not support listening for WiFi-frames in monitor mode (link in Supplemental Information).

3.3 Execution

Various tests were conducted in early stages of the project in order to get an earlier insight into which problems that could arise and potentially affect the outcome.

During the execution of these experiments, additional steps were taken in order to improve the conditions to the greatest extent possible. The smartphone was held flatly at around one meter from the ground during all of the tests. How frequently the WiFi probe broadcasts occurred was increased by manual updating of the WiFi menu within the settings of the device.

The person holding the smartphone rotated in a slow and steady manner at the five locations that were within the area, not including the positions at the edges. This was done in order to reduce potential signal blocking that could occur if the back of the subject were directed towards one of the sensors which previously have been proven to have an effect on the signal strength[2][10].

3.3.1 Mock-up

A rough mock-up of the experiment was initially set up to check the basic functionality of the system and to identify problems that could have been missed during the previous, more theoretical phase of the project. The mock-up was used to iteratively correct problems and apply solutions to problems that previously had not been thought of, that strongly could affect the results.
3.3.2 Distancing formula

When the functionality of the mock-up system had been confirmed after several iterations, the testing to find the optimal dimensions for the testing area could begin. The maximum distance between a sensor and the smartphone, where the signal would still be deemed reliable, had to be found.

To do this, the smartphone was iteratively placed 10 meters away from one of the sensors. Measurements were taken at each 10-meter point until the most optimal maximum range for our setup was found. This defined the maximum possible distance of the experimental area, and thereby laid the foundation to the size of the area where the experiments were to be conducted. This experiment also produced valuable RSSI-values which could be used for further analysis. Formula 3.1 was used for converting the gathered RSSI-values into distances of meters.

When the theories and functions had been tested and verified, a less precise version of the final experiment was conducted. The purpose of this test was to gather somewhat more accurate data. The previously mentioned mathematical formulas could be applied to this new data to calculate positions within the testing area. This provided the opportunity to fine-adjust the system in accordance with the actual scenario.

3.3.3 Final experiment

The final experiment was conducted by gathering forty RSSI values from each sensor for every location tested. The sensors were positioned freely in the air attached by the power cable to the tripods in order to avoid as much interference as possible. The mean RSSI value of the forty values was calculated for each sensor and converted into distances in meters by using formula 3.1.

The positioning formulas (function 3.2 and 3.3) required the distance from precisely three different sensors. Due to this, there were four different ways to combine the distances from the sensors into a location.

Figure 3.4: The four combinations of sensors possible for using the trilateration technique. These positions can be combined into one more accurate position.
3.3. Execution

All four combinations of sensors were processed, which resulted in four different locations. To get one unique and as precise location as possible, a mean value of the X-coordinates from all four calculated locations was derived. The same was done with the Y-coordinates. This resulted in one final position that took all of the four sensors values into consideration. Since all of the tests were done within the defined testing area, the coordinates were capped to never go outside of the testing area. This means that measurements that would influence the final location to be outside of the testing area were capped. This location was then compared to the true position of the smartphone and a deviation was calculated.

This process was repeated fourteen times at thirteen different locations. The location in the middle of the testing area was measured twice. As seen in figure 3.5, the readings were conducted in a diamond pattern for the best possible coverage of the area.

![Diagram](image)

Figure 3.5: Visual presentation of positions within the testing area that were used for testing the accuracy of the system.
Chapter 4

Results

As the distance between the sensors and the smartphone increased, the WiFi-signal strengths were observed to decay in a logarithmic manner. This behavior is presented in figure 4.1 below.

![Signal strength decay diagram](image)

Figure 4.1: Decay of signal strength when the distance from the source is incremented. The yellow line represents the strongest signal strength recorded, the red line represents the weakest signal strength recorded and the blue line represents the average signal strength recorded. 10 measurements were recorded at each ten-meter distance from the sensor. The leftmost data point is recorded at the distance of one meter from the sensor.

Since the decaying of the signal strength was logarithmic, the formula 3.1 found to convert from RSSI to meters seemed viable since it was logarithmic. The RSSI values were put into formula 3.1 to test its accuracy. The result of using this formula is presented in figure 4.2.
Figure 4.2: The result of putting the RSSI-values (previously displayed in figure 4.1) into the function 3.1. The red line represents the highest distance calculated for a single data point, the yellow line represents the shortest distance calculated for a single data point and the blue line represents the calculated distance based on the mean signal strength. The green line is added as a reference to what perfect conversion would look like.

Distinctive spikes are easily recognizable in figure 4.2. An example is the 206.5 meters calculated at the 70 meter mark with a deviation of 136.5 meters from the true distance. Even though some of the measurements significantly deviated, the mean value of all measurements still remained close to the true distance.
Figure 4.3: Deviation from the true distance of the calculated distances from figure 4.2. The red line represents the highest distance calculated for one single data point, the yellow line represents the lowest distance calculated for one single data point and the blue line represents the distance calculated from the mean signal strength.

As seen in figure 4.3, the mean calculated distance comes very close to the true distance, never deviating more than 16.7 meters, and staying at a mean deviation of 5.6 meters.

The final experiment yielded fourteen scatter plots, one for each location tested. Each plot displays the true position and the calculated position for each location tested. These positions had a mean deviation of 16.6 meters from the true positions. The root mean square error of the calculated positions was 19.5 meters. The lowest deviation was 3.4 meters and the highest deviation was 37.2 meters. These plots are presented below in figure 4.4.
Figure 4.4: The fourteen tests of thirteen different locations within the area. The red dot represents the position the smartphone was located at. The Blue dot represents the position calculated by the system based on the signal strength of probing broadcasts.
Chapter 5

Analysis

The initial tests run on the system showed that the height at which the sensors were placed was very critical. Placing the sensors close to the ground negatively impacted the sensors ability to catch the WiFi-signals. The sensors had to be raised about 90 cm above the ground to get reasonable readings. This most probably depends on the radio-waves bouncing off the ground, thus weakening the strength of the signal at very low heights. This has proven to be a common observation in experiments similar to ours, such as the one conducted by Dong Q. and Dargie W. They placed a sensor at their ankle, knee, and waist. The position was maintained, but their results varied greatly[16]. This implies that there could exist a possibility to strengthen and stabilize the signal further, using various heights.

In addition to this, nearby objects such as walls surrounding the area nearby could have an enhancing effect on the signals as they get reflected and bounce back into the area, thus resulting in a stronger signal. On the contrary, obstacles within the measurement area would instead block the signal which would raise the possibility of inconsistent or generally lower signal strengths. In our case, the area our tests were conducted in had neither obstacles nor any walls nearby. Therefore we consider this effect to be negligible.

The placement of the sensors which define the area was positioned by using a measuring tape. We estimate that the margin of error could be up to 50 centimeters in the distance between each of the sensors. This could have some slight effect on the signal strengths, but probably not enough to cause any notable differences in the positioning.

Environmental conditions as previously described would apply nearly everywhere in a more realistic scenario, hence why the formula 3.1 for converting signal strengths into distances contains an environmental variable. Its intentions are to parry such outliers in signal strengths that are caused by surrounding obstacles that almost always exist in realistic environments. This variable could be useful if potential obstacles in between the source of the signal and the sensor would be known in beforehand. These obstacles could then be accounted for, but since different obstacles give different resistance and are placed in different locations, this variable is hard to use in a reliable way. In our case there are no obstacles what so ever. That is the reason for why we chose the value 2 for the environmental variable. It corresponds to an open environment free of any obstacles that could amplify or degrade the signal.
The use of equivalent hardware with the same capabilities would, without doubt, have the possibility to affect the measurements significantly. There have been studies that clearly demonstrate this behavior between different WiFi chipsets, even though they could be manufactured by the same vendor \cite{11}. The design of the antenna could as well play a big part in the differences\cite{7}.

Figure \ref{fig:signal_strength} illustrates how volatile some individual signal strength measurements can be in regard to the expected signal strengths, distinctively seen at the 70 meter mark on the X-axis. It also illustrates how the accuracy improves as more readings within the same location are taken into account, which is why we chose to take up to 40 data points for every position tested. This is a common behavior in the area of location positioning techniques, such as the well known and highly regarded positioning using GPS\cite{8}.

This is the reason we take several readings from every sensor at the same location and thereby base the signal strength on their mean value instead. This was also done in a similar study in which they took around one hundred RSSI values, and in some cases, used the average value of them \cite{16,11}. The negative aspect of this accuracy-increasing method is of course at the expense of the time spent at each location. As our approximated spent time at each position was around 5 minutes, it would unfortunately be very impractical in a more realistic scenario.
As this project involves monitoring publicly available WiFi-signals, we deemed it appropriate to elucidate it to Blekinge Tekniska Högskola’s southeast region of the ethics committee, the county government of Blekinge and Ph.D. Martin Boldt whom has excellent knowledge about the area.

The county government reflected their point of view through the Swedish PTS (Post och Telestyrelsen) who referred to a previous project with similarities to ours. Their department of networking assessed that the Swedish law (2003:389) didn’t cover this type of activity. Although, the Swedish Data Protection Authority later concluded that constant monitoring and gathering of MAC-signals was in violation of 10§ of the Swedish Personal Data Act which is published online (link in Supplemental Information). Our project doesn’t revolve around mass collection of MAC-addresses. We also plan to filter everything based on the MAC-addresses of our own devices to which we have full consent. This leads us to the conclusion that these laws wouldn’t affect our work in regard to the ethics.

Both the representative of the southeast region of the ethics committee and Ph.D. Martin Boldt deemed that our idea is viable and not violating any one’s privacy as long as we kept to our plan of automatically filtering the data on our devices.

With these statements taken into consideration, we proceeded as planned with the project.
Chapter 7

Conclusion and Future Work

As of this study, we have presented a cost-effective positioning system that is capable of generating somewhat accurate positions. By using sensors worth less than 500 SEK in total in an area of 6400 m$^2$, we managed to position a smartphone based solely on WiFi-probe broadcasts. The average deviation in the accuracy was 16.6 meters, and it’s root mean square error resulted in 19.5 meters.

The accuracy of the system may not have been extremely high, but definitively high enough to prove a point.

With very cheap components we managed to create a system that, without the potential knowledge of the smartphone’s owner, could position the smartphone within a large area. More and better sensors would help yield a more precise position and could also cover a larger area.

To improve the functionality of this system there are two different approaches we can think of. Either to improve the quality of the readings so that fewer readings are required to derive a position, or to generate readings faster by having more sensors listening so that the same amount of measurements can be generated in less time. The more broadcasted frames received, the better accuracy can be offered.

One interesting aspect would be to compare the results that we received by using the 2.4GHz band, in contrast to utilizing the 5 GHz band instead. One could augur that the 5GHz signal would be more robust as it is free from the cluttered 2.4GHz band, but the shorter transmitting range would probably shrink the total possible measurable area slightly\cite{11}.

Everything comes at a price. Higher accuracy will take longer time to achieve. Faster positioning will yield lower accuracy\cite{8}. The ones implementing the system has to choose what is more important to them. Using higher quality hardware could minimize the trade-off between positioning speed and accuracy, but the payoff will most probably never be eliminated entirely.
References


Appendix A

Supplemental Information

Patch for Raspberry Pi Zero W to activate WiFi-monitor possibilities:

https://github.com/seemoo-lab/nexmon

We have used Airodump-ng from Aircrack-ng

https://www.aircrack-ng.org/doku.php?id=airodump-ng

Link to verdict by the Swedish Data Protection Authority

sendToServer.sh

This is the script that is being autonomously run on the sensors to report data to the server for further calculations.

#!/bin/bash

case `hostname` in
  "Node1")
    port="1111"
    ;;
  "Node2")
    port="2222"
    ;;
  "Node3")
    port="3333"
    ;;
  "Node4")
    port="4444"
    ;;
esac

# Send cellphone signal-strength to server
while [[ 1 ]]
do
  # Grab signal strength and send it to server through netcat
cat /home/pi/measuring-scripts/OUTPUT-01.csv | grep 24:92:0E:83:23 |
cut -d " " -f 6 | awk '{print substr($0,2,length($0)-2)}' |
ncat --send-only `route -n | grep UG | grep -v UGH` |
cut -f 10 -d " " $port
  # Wait for signal to update
  #Sleep is used to avoid spamming the server
  sleep 5
done
Listener.sh

This script starts a positioning session. In-parameters define how many measurements should be taken on each sensor.

```
#!/bin/bash

if [[ $1 == "stop" ]];
then
    echo -e "Stopping potential zombie listeners..."
    echo "stop" | ncat --send-only 127.0.0.1 1111 2>/dev/null
    echo "stop" | ncat --send-only 127.0.0.1 2222 2>/dev/null
    echo "stop" | ncat --send-only 127.0.0.1 3333 2>/dev/null
    echo "stop" | ncat --send-only 127.0.0.1 4444 2>/dev/null
else
    bash listener stop
    echo -e "Removing data-files..."
    rm /home/hanke/examensarbetet/server-script/.data*  
    echo -e "Removing coordinates file..."
    rm /home/hanke/examensarbetet/server-script/coordinates
    echo -e "Starting listeners..."
    bash /home/hanke/examensarbetet/server-script/.nc-to-file 1 $1 &
    bash /home/hanke/examensarbetet/server-script/.nc-to-file 2 $1 &
    bash /home/hanke/examensarbetet/server-script/.nc-to-file 3 $1 &
    bash /home/hanke/examensarbetet/server-script/.nc-to-file 4 $1 &
    echo -e "Waiting for data..."
    for count in `seq 1 4`
    do
        echo `ncat -l --recv-only 5555`
    done
    echo -e "All nodes done.\n"
fi
```
This script is started one time for every sensor from the listener-script. This script collects the data from the sensors, makes sure it is new data and saves the data to files.

```bash
#!/bin/bash

lastknown=""
readings="0"

case $1 in
    "1")
        port="1111"
        filename=".data1"
        ;;
    "2")
        port="2222"
        filename=".data2"
        ;;
    "3")
        port="3333"
        filename=".data3"
        ;;
    "4")
        port="4444"
        filename=".data4"
        ;;
esac

while [[ $readings -lt $2 ]]; do
    temp=`ncat -l --recv-only $port`
    if [[ $temp == "stop" ]]; then
        exit 0
    else
        if [[ $temp != $lastknown ]]; then
            echo "$temp" >> /home/hanke/examensarbetet/server-script/$filename
            lastknown=$temp
            readings=$((readings+1))
        fi
    fi
done

echo "Node$1 reported in done." | ncat --send-only 127.0.0.1 5555
```