An IoT-based Retail Store Using Indoor Human Localization and Predictive Analytics

Melchizedek I. Alipio

Introduction

Indoor positioning system (IPS) became a new way of locating a position of a device in a confined area. The creation of IPS filled the gap of Global Positioning System (GPS) that is only restricted to measuring a position in an outdoor setup. The use of indoor positioning system can be applied to numerous applications one of which is for retail analytics. Philips Company used this method to create a frictionless shopping experience for customers. Other supermarkets in Dubai, Netherlands, and U.S.A also adopted the same method because of its advantages to both the customer and the retailer. The advantages of using IPS in retail analytics is to give convenience to shoppers by showing them the optimal routes to take and suggestions on what products to take. Retailers can also assess the day-to-day operations using indoor positioning system as monitoring tool to see which route is frequently passed by the customers. The integration of IPS with retail analytics became possible using IoT systems.

Although indoor localization is a topic that has been around the automation process for a long time, even with all the time, cost, and efforts dedicated on this idea, it cannot be implemented perfectly to every indoor setup similar to what the GPS has achieved. The main restriction of the existing indoor localization methods is that many factors that can affect a signal are present indoors so most outdoor tracking techniques would not be applicable for indoor use. Several methods were proposed and still being developed up until this day, RFID and Camera localization method were two of those devices used for this application due to their availability and identifying capability. RFID method mainly uses passive tags for localization which is characterized by small radius of detection with no need for an internal source, these tags are
implemented due to their reduced cost. In this study, an active RFID tag would be used so as to increase the radius to be detected and increase the overall efficiency of the devices in terms of costing. The problem with this implementation is the reliance of the system to the complex algorithmic method that would be used in order to track each individual active tags on the area. The camera method on the other hand, relies on the high resolution image released by the camera in order to track the person through the indoor location using complex image processing algorithms. The camera’s advantage is that it has the capability to track people using a complex algorithm, although within a large scale implementation, positioning will be the key especially when there is no clear view of the person's position and when there is a lot of people being tracked at the same time. To the best of the authors’ knowledge, a combination of RFID and Camera implementation in an indoor localization is not yet fully evaluated and implemented to its fullest capacity.

The main idea of this work is to investigate the effects of combining the RFID and Camera implementation to an indoor localization for a more accurate tracking by mainly implementing them together in the same network. Since the RFID has accuracy problems resulting from detecting tags in a certain area enclosed due to the interference experienced by the devices, Camera implementation would be appropriate as a supplement device in order to increase the RFID tracking capability by means of integrating both of the data received. By implementing this type of localization in a market, retailers can track the position of customers on a specific aisle which can also be used for retail market strategic analysis. Also by implementing this setup, application based on the indoor localization where in data gathered from the implementation would be used as a means of giving best route inside the market as well as giving emphasis on the most purchased items and the less purchased items to increase the
overall sales of the store. IPS has been a very crucial topic since no implementation has been seen as the best localization method for large indoor setup, so if this study results in a higher tracking capability in diverse conditions, then this could be a big stepping stone to the development of a standardized IPS method.

The study aims to develop a localization technique based on RFID nodes and cameras to be used in retail analytics for market prediction. In line with the general objectives are the following:

- To develop a network of RFID nodes and camera to gather information from human activity in order to create an indoor localization system for a supermarket setup.
- To design a set of conditions and algorithms that will complement the data gathered from both RFID and camera nodes to produce an accurate localization output.
- To evaluate the accuracy of the localization implementation based on RFID and Camera nodes in a controlled area using statistical analysis.

The significance of the study aims at the application that the IPS can be applied to, such as in the case of this study, for a market retail analysis. It aims to show the significant advantage of implementing two devices rather than using a single device or method in locating relative position. This can lead to an increase on the progress made for standardization of an indoor localization techniques that has been for years, became stale due to the disadvantage seen on most common types of indoor localization methods and devices used. The study also gives an emphasis of using machine learning algorithms as prediction for market analysis in order to show the effects of applying the proposed system in a market setup. The system in turn, affects the shoppers when the market has been modeled based on the prediction and after the system has
been coupled with an application that would give users a brief piece of data that would be useful to them such as most frequently bought items, recommended path, etc.

The scope of the study covers mainly the IPS techniques such as trilateration and RSSI in order to track a person’s position in a controlled area, RFID implementation on a large indoor setup reinforced with a strategic placement of camera nodes in order to fully utilize the data that can be gathered upon implementation of the proposed system, and the tools and other techniques that constitutes on implementing both devices on a localization setup. However, limitations upon implementing the system has to be considered due to various constraints. The first limitation of the study falls on the area of implementation. Deployment of the devices is dependent on the area of localization, so the market that would be subjected to the study must be chosen in order to fit the type and number of devices to be deployed. This system is also solely designed for a market retail setup, therefore this system would not fit on any other application without major redesign for both the devices and the algorithm used. Precise tracking requires a long training time in order to calibrate, test, and get the best results to describe the appropriate capability that the device can handle, the time it takes to perform such operations varies with the type and size of the environment to be implemented. External factors that are uncontrollable such as electrical noise, ambient temperature, humidity, obstructions and other different scenarios and type of materials that can affect the reading of both devices are also considered in the limitation of this study. Lastly, the system setup is designed to detect customers carrying baskets with tracking device, so people within the market that doesn’t carry any basket or pushcart would interfere with parts of the systems design, although the system can locate and count those people, it can only do so to some extent so therefore they can still interfere with the whole process on rare cases and scenarios.
Related Works

A recent study [1] revolves around the use of Passive RFID tags and readers to determine the user location by tomographic imaging and signal strength. Passive RFID tags are cost efficient compared to Wi-Fi tags and Bluetooth beacons simply because it doesn’t need an internal source or supply, instead it relies on readers to transmit power so that it can use that same power to transmit its own identity tag back to the reader. The main advantage of using these tags is that clustered placement can be cost efficient, one can place multiple tags in a small area for accurate detection without suffering much in terms of cost. This is true for this approach, since the tomographic imaging based on RFID centers around the passive RFID tags transmitting low powered signal to determine and map the location, one must place multiple tags on a certain area for accurate response. Similar to a camera approach, the area being localized is divided into segments containing a specific passive RFID tags, these tags will transmit a signal directly to the RFID reader.

Using the attenuated signal and knowledge on the RFID tags unique location; the reader can determine the attenuated signals exact location that therefore contributes to the approximate location of the person being tracked. This approach can be as precise as that of a complex camera localization approach. However, this implementation must consider the fact that the RFID reader cannot differentiate the number of people contained in a specific dimension when the signal is thrown back, also the RFID signals can strongly be affected by environmental objects such as metals which is present in most indoor location. This approach must take into account the environment and the continuous power consumption that is used due to the continuous transmission and reception of signals brought about by implementing multiple passive RFID tags.
and readers. The article considered the strategic placements of each passive tags in order to utilize the unique location of each tags from the given environment.

As mentioned above, camera has been one of the more accurate approach for indoor localization, this is because a camera can be able detect an object and process that detection real time. One example of camera usage in localization is the paper written in [2] where they proposed to use a panoramic camera in order to detect and localize an indoor location. The process was divided into three parts; the first part is the image processing where in the camera captures the location that one wants to localize. The image captured by the camera is resized and rotated in order to obtain a more accurate localization per pixel. Then that image is converted to a gray one in order to reduce the calculation complexity for a multi colored image. The second part is to calculate the background image by comparing the previous and current image of that same location using a complex computational scheme. The last part involves computing for the coordinate location based on the foreground pixel reference in the center of the camera. The result achieved a mean error of 0.37 m which is a lot smaller than that of other technique such as the trilateration and fingerprinting method.

The work provided a complex and accurate means of localization but the limitations can be prevalent for this type of method. One aspect is that the camera can only process or locate objects and people close to the center of the camera used for pixel reference, also the panoramic camera is not only costly, but it also has a limited range capability depending on the height of the placed location. So increasing the overall area to be localized can contribute to the error produced by the camera due to objects being farther away from the pixel reference point which can be out of range for the camera’s span. Therefore, the farther away the person is from the
reference center node which can occur due to the increase in the demanded area to be localized, the less accurate the detection will be.

Lastly, the work done in [3] on Camera-RFID localization system is an important work that complements the author’s study due to the fact that their work showed the feasibility that both of those devices could work as one localization method with accurate results. They implemented an Unscented Kalman Filter to complement the RSSI values from the RFID in order to determine the exact location of the object being tracked. The main point of their study is that they will use the camera as a means of detecting if the object is present within the trolley, while the RFID tracks the location of that said trolley. In simple terms the camera acts as a confirmation that the object being track is indeed equipped with a passive RFID and should be tracked by the algorithm.

This is the reason why a low quality camera was used because precise pixel detection was not needed and long range detection using camera was practically non-existent. They focused on converting the Kalman Filter into a Constrained Kalman Filter in order to fit it to that specific application. The problem with this type of setup is that the camera used was underutilized due to the fact that it doesn’t necessarily locate the object but it only determines if the object is present which is quite unnecessary as it can be done by a separate RFID network. This work mainly prioritizes localization using RFID which makes the Camera-RFID localization method still under used.

Based on the works presented, most indoor localization utilizing RFID or camera suffers between two major problems; a larger indoor area with multiple objects, and Data misinterpretation which can both contribute to inaccuracy of tracking. Both RFID and camera have problems when localizing a large area. For camera, the problem is that the larger the area,
the more the camera reading suffers due to long draw distance present in the images, also objects that blocks the view of the camera would contribute to problems such as camera misinterpretation. To compensate for this, multiple cameras must be placed around the whole area and each must be able to communicate with one another and be able to identify a person, therefore cost and complexity quickly arises. This same problem goes for RFID, as the area increases, the number of either passive or active RFID tags and RFID readers must be increased to track the person properly indoor, also since RFID is signal dependent, an object present between the reader and the tag can contribute to the overall inaccuracy of reading due to attenuation. Many factors can be considered to increase the accuracy of RFID indoor localization. However, if one improves those factor (positioning technique), the RFID accuracy still won’t be able to surpass a camera’s reading when identifying a person due to the RFID relying solely to the signal present on the system. Based on these problems, both device’s accuracy can be theoretically increased by using strategic deployment so that both device can cover larger and more meaningful location on the indoor setup where in compensation for obstacles and objects present are prioritized. Implementation for both of these devices on the same indoor network will prove beneficial for tracking purposes. Both of the readings gathered will complement each of the localization method and therefore, can decrease the localization error for indoor tracking. If implemented successfully, indoor localization using these devices would be considered as one of the reliable method for locating relative position in an indoor set up and can be further used for applications such as retail analysis using a machine learning algorithm that will generate a prediction which in effect, will benefit to both the retailer and the customer.
Overview of IoT-based Retail Market

<table>
<thead>
<tr>
<th>Input:</th>
<th>Process:</th>
<th>Output:</th>
</tr>
</thead>
</table>
| 1.) RFID  
2.) Camera | 1.) RSSI  
2.) Trilateration  
3.) Image Processing | 1.) Foot Traffic  
2.) Frequent Path  
3.) User Interface for Retailers |

Figure 4: Conceptual framework

Figure 4 specifies the conceptual framework of this study. The framework shows the inputs, outputs and the process that will be used for this implementation. The input value comes from the RFID which will identify various parameters from the tags, specifically the parameters such as the unique ID code for each active RFID tags and the I/O parameters in terms of power. The camera input is a continuous feed of image where in its usefulness occurs whenever the RFID wants to check the conditions. The process involves the RSSI and Trilateration for the RFID localization method and the Image processing technique for the Camera localization implementation. Based on the processed data, a foot traffic for the whole market will be created, with this, the most and less frequent path will be identified and a user-friendly interface for the retailer will be generated for monitoring purposes.

Figure 5: System overview
A system overview of the proposed system is shown in Figure 5 where it is composed of multiple devices like RFID and camera that is capable of locating a relative position of a customer in a supermarket setup. The processing of the data gathered is done using a computer in order to produce a heat map of foot traffic generated. The data will be saved in a storage database on the computer so that further and future data analysis can be done. Lastly, the generation of heat maps is governed by series of algorithms and conditions which will be discussed in the following section of the study.

**RFID and Camera Based Localization Implementation**

The first method covers the implementation of a hybrid network composed of RFID nodes and camera that is capable of communicating with the base station for further processing. The RFID reader must be compatible with the RFID tag that will be used to ensure that connectivity would not be a problem. Also, in order to use Trilateration method for position modeling, there must be at least three nodes that detect the position of the object to be localized so the 3 RFID readers must enclose the same range of area. Each RFID tag has its own unique identification; this will be helpful in determining the distance of each tag received by the RFID reader. Each time the reader receives data from the tag, the reader transmits those data to the computer, the computer then organizes the data so that the distance of one tag is listed for the three reader placed. The Trilateration is done by the computer based on the signal strength transmitted by the three tags which will be elaborated further in the next method. The modeled position of each present tag is displayed on the computer using MATLAB software, these position changes per second as the tag moves and transmits different signal values. Figure 3.3 that shows the sample connection for the devices to be used for the RFID localization.
As shown in Figure 6, three RFID readers are placed on different positions and the three RFID readers are connected to the computer through a communication medium to transmit the data gathered from the three tags.

On the other hand, an implementation of camera is used to compensate the inevitable errors that may occur upon gathering the relative position of the tag using RFID nodes. The addition of camera is essential in order to make the prediction and measurements more accurate. Factors are to be considered upon setting up a camera for a supermarket set up; a) quality of the camera – different cameras have different ratings, the quality of the images that it captures vary, that is why choosing the most appropriate camera for the image processing technique to be used inside a supermarket pinpoints to a crucial decision making; b) positioning of the camera – the accuracy of the detection does not just rely on the quality of the camera that it may capture, but also relies in the placements of the camera as well. The placements of the camera must consider several conditions such as the width of the aisle, the distance between aisles, the opaqueness of blocks, and also the different height and sizes of each person that is being detected by the
camera; c) algorithm for detection – the algorithm that is being used for detection is also considered. Algorithms to be used for camera are to be elaborated more in the second method.

![Camera Block Diagram](image1)

**Figure 7:** Block diagram for camera prediction

Figure 7 shows the general process of using camera to detect the number of persons for specific aisle. The processed output is saved in the computer for further data processing. A 3-dimensional perspective setup of camera and RFID nodes are shown in Figure 8. The camera must be placed in an elevated height to completely oversee the number of people present in the aisle. The RFID readers on the other hand have to be placed in a triangular position so that the radiation pattern for each of the reader will cover the entire aisle.

![RFID and Camera Network](image2)

**Figure 8:** Perspective view set up of RFID and camera network
Hybrid Localization Algorithm

Figure 9 shows the data flow in order to localize using RFID. The main concept is the combination of the RSSI values of the three RFID reader to perform Trilateration in order to predict a model for positioning.

The trilateration process using RSSI and RFID localization method is shown in Figure 10. This method utilizes the three position of the RFID reader in order to triangulate the exact position of the tags in the area. The figure shows the intersection created by the three RFID reader in order to detect the tag, this method is done for all the tags present in the area. Trilateration has two main requirements, the 1<sup>st</sup> one is that the detecting device must enclose within its range of detection the unknown device, this is done by strategic placements, the 2<sup>nd</sup> requirement is that the distance between the detecting device and unknown device to be detected must be known, in order for this to be achieved, RSSI function must be used.
The measurement of the received signal between the user point tag and the three readers are done mathematically using Friis equation which describes the RSSI function. The formula shows the Friis equation where in the variables $P_t$ is the transmitted power by the RFID reader which is provided in its specification sheet. $P_r$ is the received and transmitted power by the tag and it is measured using the reader itself. The wavelength can be calculated using the ratio of speed of light and the center frequency of the RFID reader stated in the specification sheet.

The R value in equation states the approximated distance between the reader and the specified tag in meters. However, attenuation is a factor to be considered as present due to walls and other external factors. The list of materials that can be causes of attenuation for a 2.4 GHz RF wave are listed on Table 1. Considering that this study is in the case of a supermarket set up, attenuation of materials like steel frames, concrete walls that interfere with the received signal
between the reader and tags have to be considered in the mathematical computation in order to minimize the error of the measured result.

Table 1: Attenuation of Different Materials (Acquired from [4])

<table>
<thead>
<tr>
<th>Material</th>
<th>Attenuation (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.4 GHz</td>
</tr>
<tr>
<td>Interior drywall</td>
<td>3-4</td>
</tr>
<tr>
<td>Cubicle wall</td>
<td>2-5</td>
</tr>
<tr>
<td>Wood door (hollow – solid)</td>
<td>3-4</td>
</tr>
<tr>
<td>Brick/concrete wall</td>
<td>6-18</td>
</tr>
<tr>
<td>Glass/window (not tinted)</td>
<td>2-3</td>
</tr>
<tr>
<td>Double-pane coated glass</td>
<td>13</td>
</tr>
<tr>
<td>Bullet-proof glass</td>
<td>10</td>
</tr>
<tr>
<td>Steel/fire exit door</td>
<td>13-19</td>
</tr>
</tbody>
</table>

The acquired distance from the Friis equation will be used in order to determine the unknown value of x and y which corresponds to the tag’s coordinates; wherein these values will be substituted into the system of equations that is shown below, in order to implement the trilateration process.

\[
d_1^2 = (X_1 - X)^2 + (Y_1 - Y)^2
\]

\[
d_2^2 = (X_2 - X)^2 + (Y_2 - Y)^2
\]

\[
d_3^2 = (X_3 - X)^2 + (Y_3 - Y)^2
\]

Variables \(X_1, X_2, X_3, Y_1, Y_2,\) and \(Y_3\) from equation 2 corresponds to the distance measured by the Friis equation based on the signal strength transmitted back to the 1st, 2nd, and 3rd reader respectively; wherein these values are the known position of the readers in a Cartesian plane system. X and Y can now be solved using the three equations by means of manipulating variables and algebraic process.
A method called sectioning will be used in order to gather position of the customer who passes in a specific aisle. The total area allotted for gathering the foot traffic will be divided into a series of boxes which corresponds as a supplementary method for both camera and RFID nodes. The boxes will be measured manually using the RFID measurements in order to determine the range of values located per section. Boxes are also labeled and grouped based on what aisle it belongs to. As shown in Figure 11 – whenever a customer steps on to a specific box, the RFID device will be notified that there is a presence of customer/s within that section. The number of tag/s received by the RFID devices will now be fed to a camera which is also located to the similar aisle where the presence of tag is detected in order to checked if there are customers who were not detected by the RFID device, the camera would look for specifically customers who do not have RFID tagged basket. The detection and checking process of the Camera will be based on the number of heads it detected within the aisle of coverage which will be expounded further in the proceeding discussion. The main job of sectioning is to be able to
still locate precisely even if minimal attenuation occurs as well as to be able to help the system for the whole localization.

The flow diagram shown on Figure 12 is the combination of RFID and Camera in order to localize. The main function of the camera is to check whether the RFID localization value correctly detected the person or the RFID became subjected to attenuation which in turn generated a wrong calculation of position. On the flow diagram, once the tag enters a certain threshold position (this threshold corresponds to the market itself) it means that a person entered a tracking range inside the market. The coordinate updates and loops in order to give the position every cycle inside the market. After initialization of values and calculation of the RSSI as well as trilateration, several loop process are done. For this code, only tag 1 was processed, this implies that the code for the whole process would be longer since more tags will be involved as well as more aisles and therefore more grids.

Each grid will be checked; so different threshold grid values will be tested for the whole area, once the grid value has been found the camera process would take place. Using the circle detection, the camera would detect if there is a circle positioned at the grid selected, if there is therefore the person was tagged and marked correctly, however if no circle was detected, the RFID generated a wrong value which therefore the camera detection must check adjacent sections in order to verify the existence of the person using that specific tag. By also using this type of detection, people without baskets and therefore no tags can be calculated and tallied for market analysis output.
Hybrid Localization in Retail Analytics

The relative position of the customers in a supermarket setup gathered by the hybrid localization system will be capable of providing applications in such a way that retailers can be able to assess its market sales. Using unsupervised type of machine learning under classification algorithms guided by retail analysis, the gathered data by both RFID and camera can furthermore be used in generating a model that will provide additional basis for evaluating market sales such as foot traffic, total dwell time of customers per aisle, the least and most frequent path used by the customer, and the busiest time of store operations. The overall system is shown in Figure 13
wherein retailers can be able to access the interface using a web browser. For this type of application, a cloud based system would fit the best since implementing a cloud system would allow retailer to access the data created by the system using devices as long as it is connected to the internet. Figure 13 also shows a sample of the heat map part of the web interface that would be shown whenever the retailer wishes to access the system using a processing unit or a mobile phone. This web interface would be programmed to provide a user friendly interface for customers as well as providing the ability of the retailers to evaluate and observe different behaviors present in the market such as the number of people per aisle on a specific time, total number of people present on the market and applications related to retail analytics.

![Smart retail analytics system overview](image)

**Figure 13:** Smart retail analytics system overview

**Deployment and Test Evaluation Progress**

This work is on the process of deployment of RFID and camera set-up in an actual retail store in the Manila, Philippines as part of data gathering procedure. It will be conducted for 30 consecutive days. Software and user interface are on the process of development as well as the cloud server set-up.
References


