Name	Ooi Jing Zhi
Email Address	ooijingzhi@gmail.com
Title	Proposing a Solution Using the Internet of Things to Understand Shopper Foot Traffic in Retail Chain Stores

Table of Contents

Introduction

Foot traffic – the number of customers in an establishment – is an important metric for retail stores to keep an eye on, since while high foot traffic does not necessarily lead to high sales, high sales is definitely observed alongside high foot traffic. With the rise of highly efficient ubiquitous sensor networks and an inversely proportional size to computing power of such devices, it can be detrimental for such business entities to obtain information about the customers in their buildings to understand their behaviour in order to construct business plans and directions to increase revenue. In this document, a solution is being proposed to utilize the Internet of Things to collect this information in varying granularity, by taking advantage of the most personal device that a customer can carry.

Problem Statement

The foot traffic of customers in a retail store can be inferenced through a variety of methods -- the most popular one in the 20th century is through the use of purchasing history of customers as the source of data analytics. In the early 2010s, Target, a major retail store chain in the US, developed an algorithm that is capable of deducing if a customer is pregnant through their shopping information down to the specific week, and famously was able to create targetted ads based on their pregnancy score before the knowledge of the new life was made known to the family members of the customer.

The issue with this method is that, it offers non-significant granularity to the owner of the building and recipient of the data. The basics of the method cover identifying shopping patterns of customers through purchase history, but is difficult to pinpoint the use case of a specific customer without a common identification like a membership card. The data also does not reflect the time spent by the customer in the individual areas to produce an efficient data set for user behaviour identification.

The results of the proposed solution should offer a practical method of obtaining a high granularity regarding the foot traffic information of customers in a retail store concept.

Proposed Solution

## Summary

The proposed solution takes advantage of the average default phone setting on modern mobile phones, which is that the Bluetooth antenna on the smartphone is always active, unless the user explicitly switches it off. Every entity with Bluetooth functionality is entitled to a unique MAC address. Using the assumption that every shopper holds at least a device with Bluetooth capabilities, it can be inferenced that shoppers carry unique devices that can be used to recognize them (since MAC addresses are constant until reset, returning customers can also be recognized without the need of additional procedures or tools, e.g. membership cards/credit cards)

To carry out this implementation, Bluetooth beacons are required. These beacons are small, individual devices that possess Bluetooth capabilities -- they are able to broadcast out a connection request and receive signals from devices that responds to the request, no matter if the connection is refused by the target device. These beacons possess little to no processing functionalities, and upon obtaining the information that is sent out by nearby Bluetooth devices, they send said information to a central warehouse through a wireless connection to be logged.

We have established that beacons are able to locate a customer through reacting when a device with Bluetooth functionalities (on/off irrelevant) is in range. To go one step further: among the information obtained through pinging the target devices, signal strength is one of them. Through comparing the signal strength between two beacons in known locations receiving the same signal, the relative positioning of the customer can be accurately tracked in the building, and the accuracy increases with the amount of beacons receiving the same signal.

Infrastructure Diagram
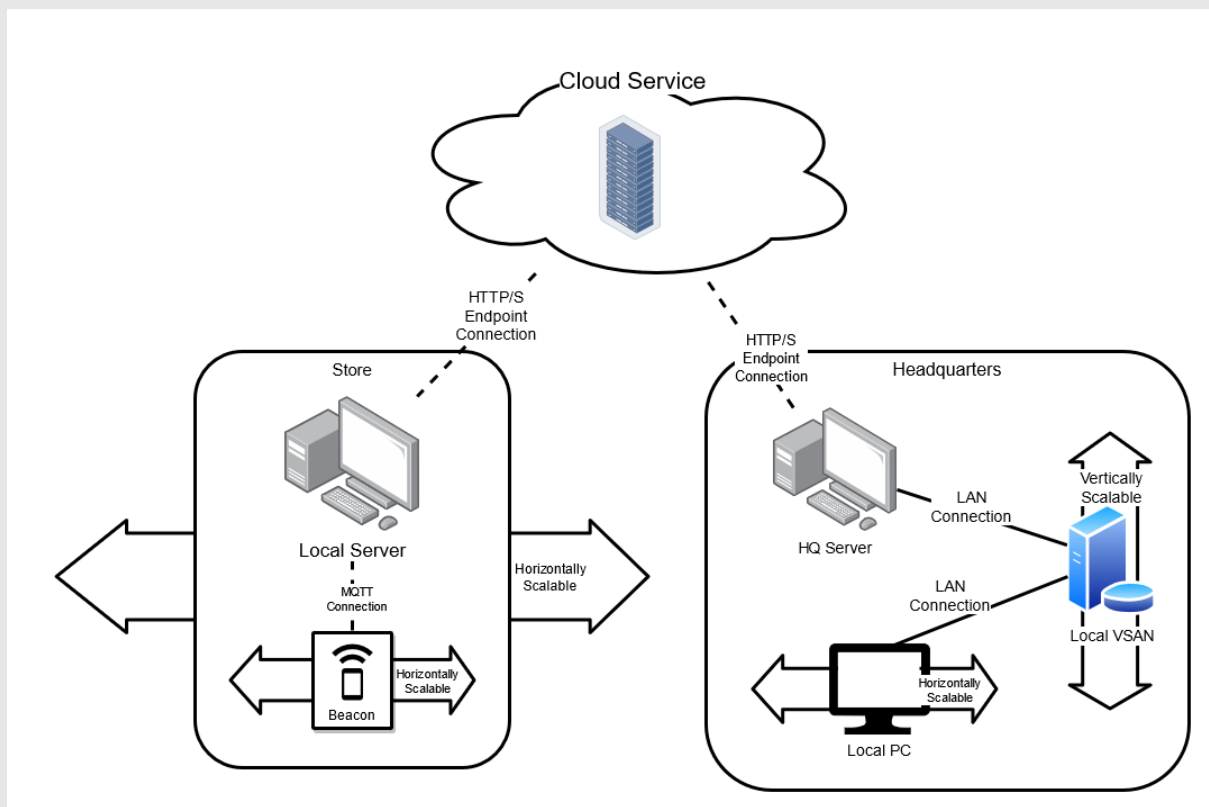
Strengths

- A high granularity of customer behaviour can be obtained, including:
  - The duration that the customer spent in specific areas/aisles of the store.
  - The browsing patterns of the customer inside the store -- even if the customer is not planning to buy anything from the area or at all.
  - Recurring customers can be identified accurately as long as they carry the same device.
- The collected data is highly versatile.
  - The information offered by the collected data can be paired with purchasing history data analysis to produce insights about the shopping:browsing habits of the customer.
- The implementation is designed to be horizontally scalable and easily manipulated.
  - Through wireless connections between beacons, they can be moved around in the store to adhere to layout changes, and added/removed from the network without much hassle.

Weaknesses

- Energy consumption
  - The wireless feature implies the usage of energy storing components (batteries) to be used in the beacons. While Bluetooth Low Energy (BLE) options are available, there may exist a need for the batteries of the devices to be changed occasionally (depending on the interval of output and capacity of the battery). A wired option can eliminate this problem, but decrease scalability and ease of change.

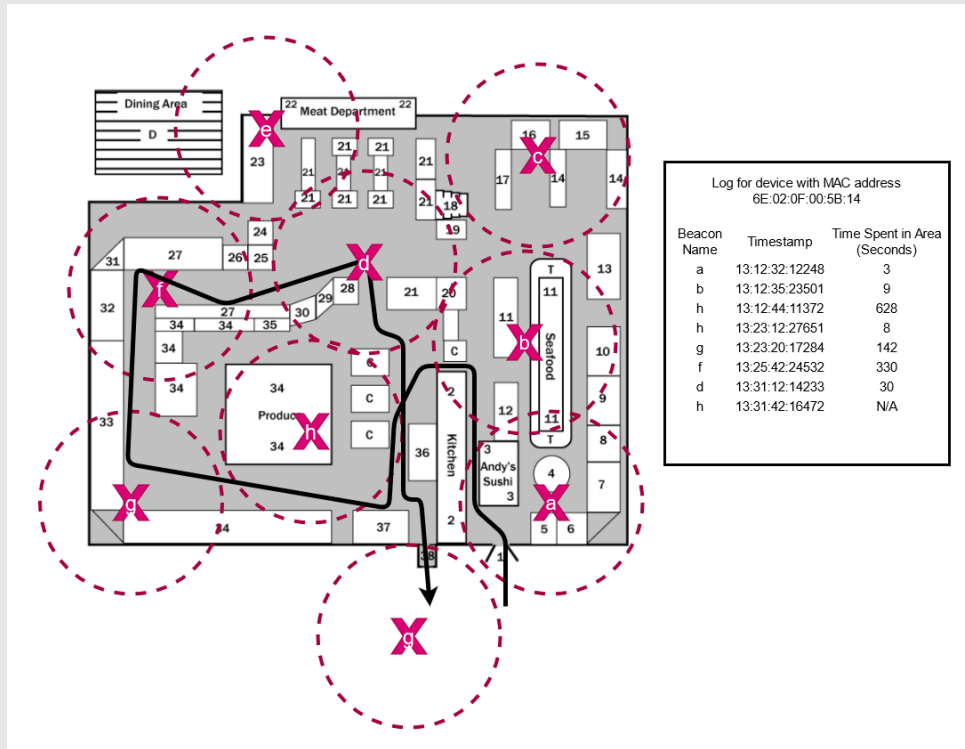## Implementation Details

1. Beacons that can broadcast and receive data through Bluetooth technology are installed in various parts of the building in a distributed fashion.
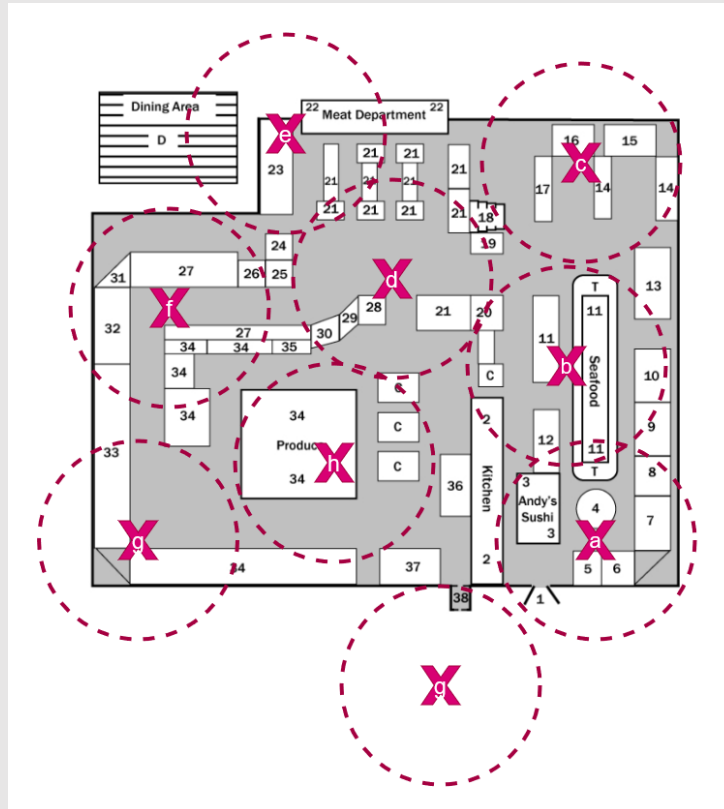
2. When a customer holding a Bluetooth-enabled device enters the range of the beacon, the MAC address of the device is timestamped and logged, and subsequently sent to the local data warehouse of the building wirelessly.

The image shows a store floor plan with various numbered sections, beacon locations marked with X (labeled a through h), dashed circles indicating beacon coverage areas, and a log table.

Log for device with MAC address 6E:02:0F:00:5B:14

Beacon Name	Timestamp	Time Spent in Area (Seconds)
a	13:12:32:12248	3
b	13:12:35:23501	9
h	13:12:44:11372	628
h	13:23:12:27651	8
g	13:23:20:17284	142
f	13:25:42:24532	330
d	13:31:12:14233	30
h	13:31:42:16472	N/A

3. The collected data is submitted to an aggregated data warehouse of the retail store chain at intervals.

4. On the store level, edge computing is used to perform rudimentary levels of user behaviour analysis to make decisions such as store layout and stock arrangement.

5. On the office level, the aggregated data is used to gain deeper insight to the overall user behaviour and trends, and perform predictions/carry out promotions at suitable times.

Privacy and Legal Concerns

According to the Personal Data Protection Act (PDPA) 2010, Act (709) under Malaysian Law, "personal data" refers to information that can be used to identify a specific individual or other information from the individual:

"personal data" means any information in respect of commercial transactions, which—

(a) is being processed wholly or partly by means of equipment operating automatically in response to instructions given for that purpose;

(b) is recorded with the intention that it should wholly or partly be processed by means of such equipment; or

(c) is recorded as part of a relevant filing system or withthe intention that it should form part of a relevant filing system,

that relates directly or indirectly to a data subject, who is identified or identifiable from that information or from that and other information in the possession of a data user, including any sensitive personal data and expression of opinion about the data subject;

However, the proposed system works on a basis in which only the MAC addresses of the customers are collected and used to label the customers in the context of carrying out activities in the building. <u>The collected data can not be used to identify the identity of the customer, and thus does not violate the PDPA using the defined boundaries of personal data and the protection thereof.</u>

Although, the Personal Data Protection Principles (Section 5) of the PDPA may be used as a guideline for data protection within the proposed solution.

Cost, Benefit, and Return on Investment Concerns

This section assumes that 1 retail stores in a district is chosen to be the pilot area over a period of 6 months. The size of the store requires them to install 15 beacons.

Item	Unit Cost	Total Units	Total Cost
BLE Beacon	RM 50	15	RM 750.00
Maintenance - Battery Change	RM 1.20 (RM 1.00 wholesale)	15 beacons x 2 batteries per unit x 2 3-month period = 60	RM 60.00
Software Licensing	RM 1,500 per month (does not increase with store numbers)	6 months	RM 9,000.00
Cloud Service Costs	RM 600 per month (120GB disk memory)	6 months	RM 0 (reusing available service)
Overhead - Site Installation - Programming • In-house	RM 4,000 per engineer per month	2 engineers over 1 month	RM 8,000
			RM 17,810

Every additional store implemented in the 6-month period incurs the following cost:

Item	Unit Cost	Total Units	Total Cost
BLE Beacon	RM 50	15	RM 750.00
Maintenance - Battery Change	RM 1.20 (RM 1.00 wholesale)	15 beacons x 2 batteries per unit x 2 3-month period = 60	RM 60.00

Software Licensing	RM 1,500 per month (does not increase with store numbers)	6 months	RM 0
Cloud Service Costs	RM 600 per month (120GB disk memory)	6 months	RM 0 (reusing available service)
Overhead - Site Installation - Programming • In-house	RM 4,000 per engineer per month	2 engineers over 1 month	RM 8,000
			RM 8,810

The benefits of the system are mostly intangible and non-immediate, and they involve being able to generate insight about the foot traffic of the customer in the store. The insights may include:

- The roadmap that the customer took in the store.
- The time that the customer spent in each section.
- Combining the above with purchase information (or the lack thereof), finding out the browsing to shopping ratio of the customer in the store.
- The reasons that caused the above.
- Product popularity in correlation to its positioning in the store.
- Product popularity in correlation to customer profile and path taken.
- Product placement that adheres to the above to generate revenue/increase accessibility.