ISE 534 - Kiana Project

Phase I
Team Master Consultants

Table of Contents

Introduction

- The Kiana data has information about time, location, ID;
  - we plan on exploring this in a way that will lead us to different business solutions
- We have researched what other RTLS companies do with their collected data, what kind of clients they have, and what type of solutions they offer
- Looking at gaps in the market and overall growth opportunities, we decided to focus on using RTLS to provide future clients with geofencing solutions

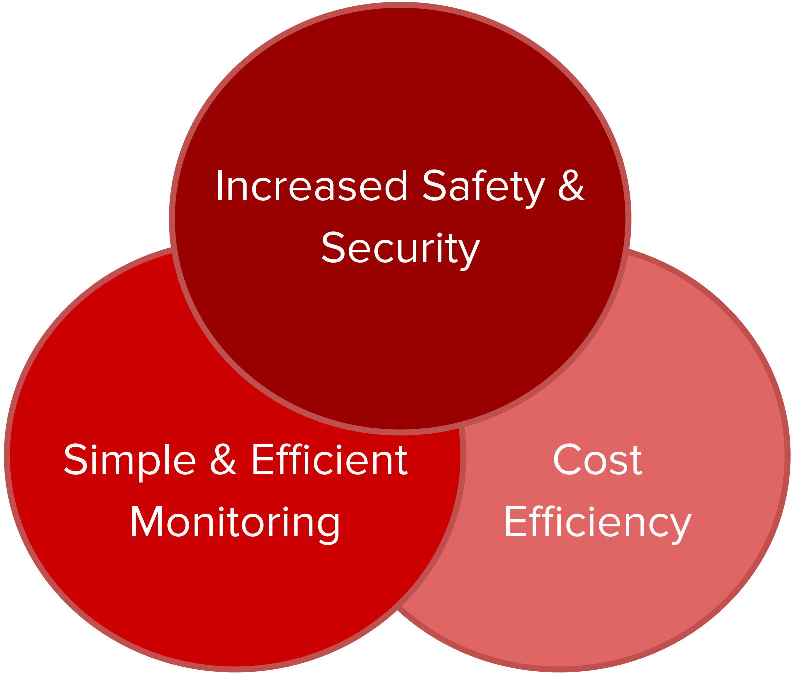
Problem Definition

- Managing security in large facilities such as factories can be difficult (especially machinery or room/building access)
- Traditional security measures such as hiring guards, enforcing badge scanning or having door readers can be costly
- It is not easy to monitor or learn where someone is/was during a certain period of time by using only security cameras or guards. Further means are often required to ensure accuracy.
- The more time spent to solve an issue in a machine, the more financially detrimental it is for the factory.

Solution Approach

- Track people's function within the facility
  - visitor, engineer, technician, security guard, etc
- Categorize the devices into fixed or mobile
- Divide the area into zones and identify its security level
- E-permitting
  - To enable/block access to certain places or use certain machines
- Virtual geofencing
  - Notifying a certain technician if they are nearest to the machine that is broken
  - If there are people in restricted locations in the building, etc…

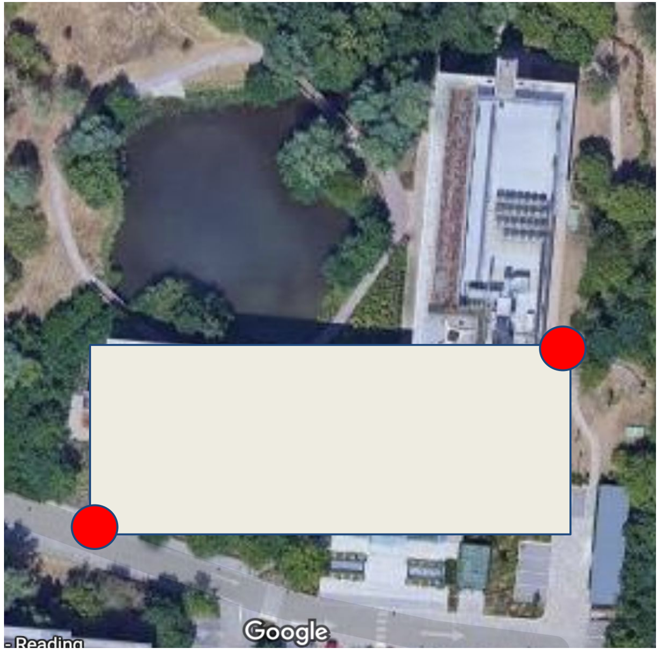
Value Proposition



Initial Data Cleaning

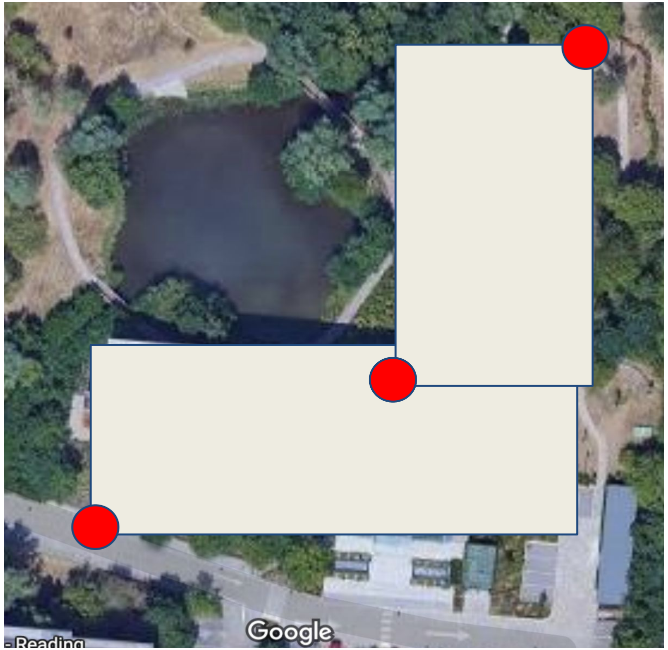
- Actual building of the data
- Remove any data points that are outside the building

Initial Data Cleaning (con't)



- Split the building into two blocks
- Identify bottom left and upper right of each block
- For the 4th floor, we only allow the first block

Initial Data Cleaning (con't)



- Split the building into two blocks
- Identify bottom-left and upper-right of each block
- On the overlapping areas, merge two blocks by removing all intersections

Initial Data Manipulation

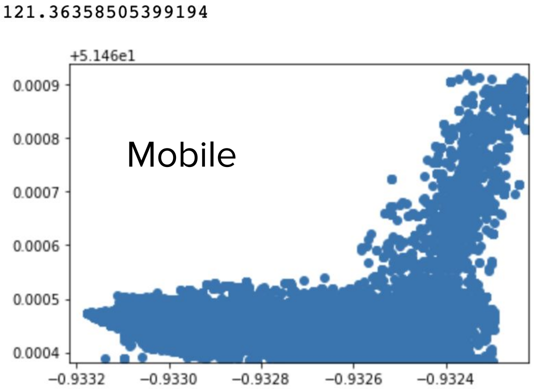
```
[3]: df_0 = read_file('data/', percent=0.2) # 16 files
data/uk_obs_coordinates_000000000040
data/uk_obs_coordinates_000000000030
data/uk_obs_coordinates_000000000037
data/uk_obs_coordinates_000000000016
data/uk_obs_coordinates_000000000076
data/uk_obs_coordinates_000000000083
data/uk_obs_coordinates_000000000055
data/uk_obs_coordinates_000000000035
data/uk_obs_coordinates_000000000061
data/uk_obs_coordinates_000000000045
data/uk_obs_coordinates_000000000050
data/uk_obs_coordinates_000000000042
data/uk_obs_coordinates_000000000062
data/uk_obs_coordinates_000000000069
data/uk_obs_coordinates_000000000004
data/uk_obs_coordinates_000000000011
```

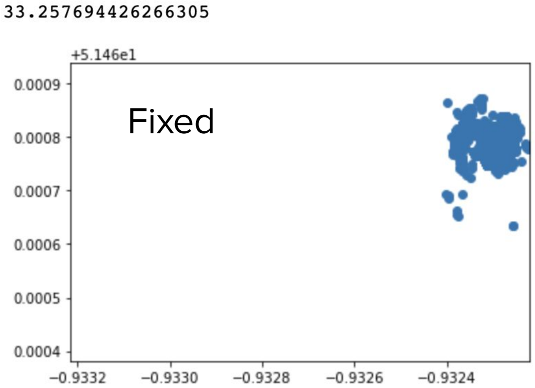
Started by using a random selection containing 20% of the data (16 sample files) to start defining rules and observing MAC Address patterns

Initial Data Manipulation

Assumptions and Rules:

- Categorize “fixed devices” by signals that ranged within an Euclidean distance < 45 meters, frequency in days > 5 days, and signals created within weekends.
- The rest of the addresses were defined as “mobile devices”





Observation and Findings

We introduced 5 new features including distance in meter, total day, total signal, weekday, and weekend to help us further classify the fixed and mobile device.

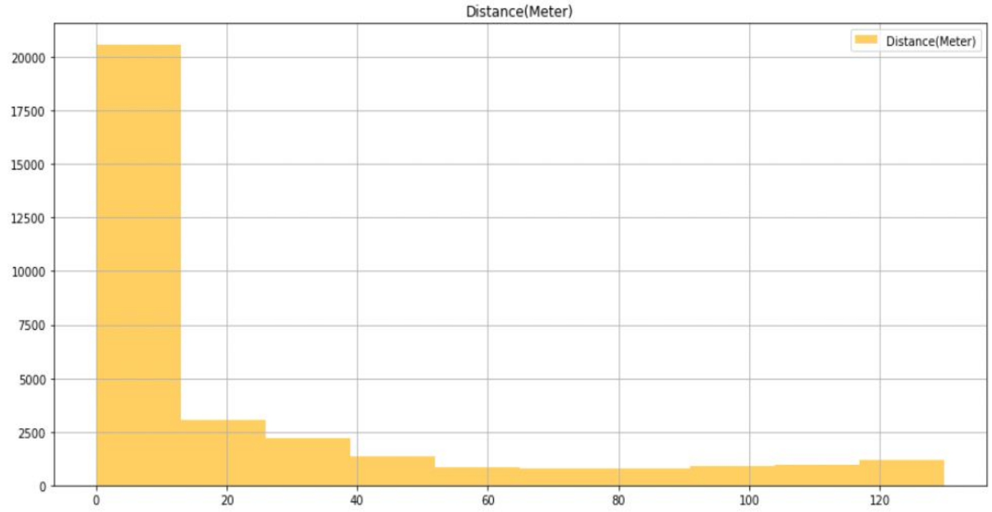
```
[64]: df_temp.describe()
```

[64]:

	max_lat	min_lat	max_lng	min_lng	Distance(Meter)	total_day	total_signal	Weekday	Weekend
count	32625.000000	32625.000000	32625.000000	32625.000000	32625.000000	32625.000000	32625.000000	32625.000000	32625.000000
mean	51.460737	51.460624	-0.932404	-0.932557	22.296224	2.588261	1393.963862	2.563648	0.024613
std	0.000173	0.000178	0.000231	0.000311	35.418779	5.498479	10340.316550	5.330616	0.527483
min	51.460382	51.460381	-0.933231	-0.933236	0.000000	1.000000	2.000000	0.000000	0.000000
25%	51.460598	51.460449	-0.932396	-0.932794	0.000000	1.000000	13.000000	1.000000	0.000000
50%	51.460786	51.460626	-0.932347	-0.932396	1.148873	1.000000	13.000000	1.000000	0.000000
75%	51.460900	51.460786	-0.932243	-0.932362	30.679037	2.000000	27.000000	2.000000	0.000000
max	51.460960	51.460960	-0.932221	-0.932221	129.850623	162.000000	605589.000000	116.000000	46.000000

Observation and Findings

Most of our data shows a distance of signals less than 30 meters.



	Distance(Meter)
count	32625.000000
mean	22.296224
std	35.418779
min	0.000000
25%	0.000000
50%	1.148873
75%	30.679037
max	129.850623

Observation and Findings

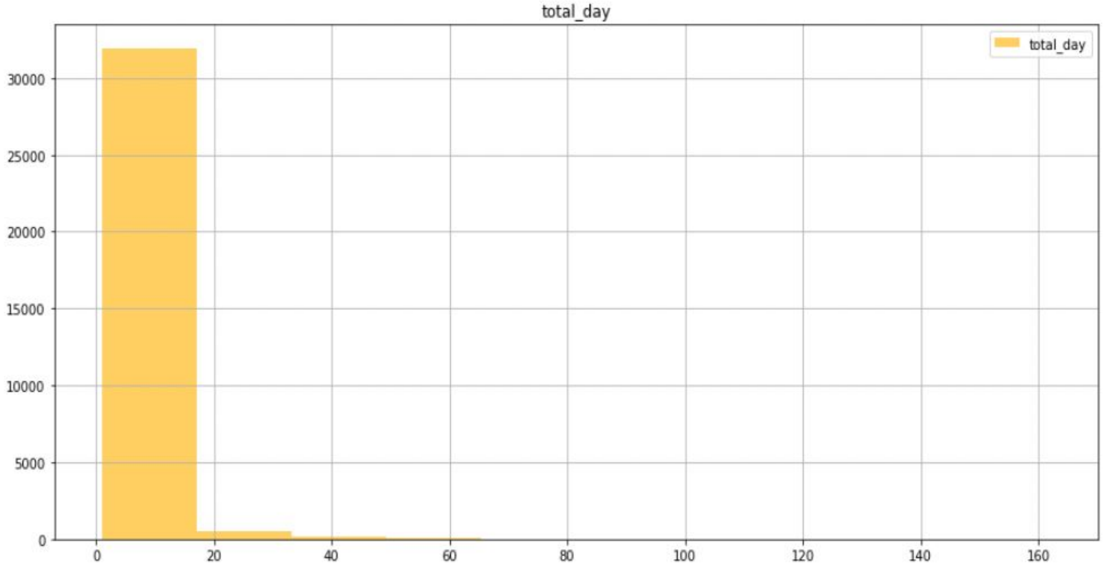
Based on our definition of having a distance < 45 meters and frequency of > 5 days , we get <u>32612 mobile</u> devices and <u>13 fixed</u> devices.

```
[73]: df_temp['Job function'].value_counts()

[73]: Mobile Device    32612
      Fixed Device        13
      Name: Job function, dtype: int64
```

Observation and Findings

Additionally, more than 75% of our data have signals in only 3 separate days, which quantifies most of our MAC Addresses as visitors.



	total_day
count	32625.000000
mean	2.588261
std	5.498479
min	1.000000
25%	1.000000
50%	1.000000
75%	2.000000
max	162.000000

Next Steps

Define Rules for Mobile Devices	Defining Building Sections	Define Associations
Define patterns to identify what mobile devices correspond to which asset type (visitor, employee and profession, etc.)	Convert the location to an industrial facility and assign different region parameters with different longitude and latitude ranges.	Define associations between mac address to find out if a person has the authority to use a machine or if the MAC address can access particular areas in a manufacturing facility

Thank you!
Questions?