INDOOR MAPPING USING WI-FI

Karan Sanwal

Shweta Sood

Sonali Garg

WHAT HAVE WE ACHIEVED ?

- **An indoor mapping model based on signals as sensitive and volatile as Wi-Fi and is something that is still invariant to the noise and rotation.**

- Dynamic and self-improving add-ons for the scalability of an already excellent architecture.

- ***A Research-Project-turned-Product. Deployed a client-server based Android Application for the project to be put in actual use.***

WHY WI-FI ?

- **Easily Available.**
- **No special hardware installations like in RFID based systems.**
- **Plenty of signals within the same building – more dimensions to discriminate.**
- **Penetrates through walls and other obstacles.**
- **Easily detected using smartphones – which are available in plenty.**

- ***However, for an effective and reliable use, that goes beyond proving the product's excellence under suitable conditions, using Wi-Fi seems to put forward a lot of challenges.***

Challenges

CHALLENGES

- ***1) Multiple sources of reflection – so different values of signal strength at the same location.***

- ***2)Wi-Fi signals degrade with environmental conditions***

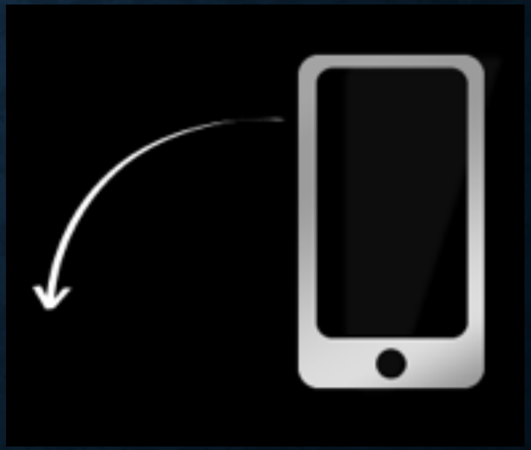
- ***3)Wi-Fi Strengths vary between different models of smart-phones***

CHALLENGES

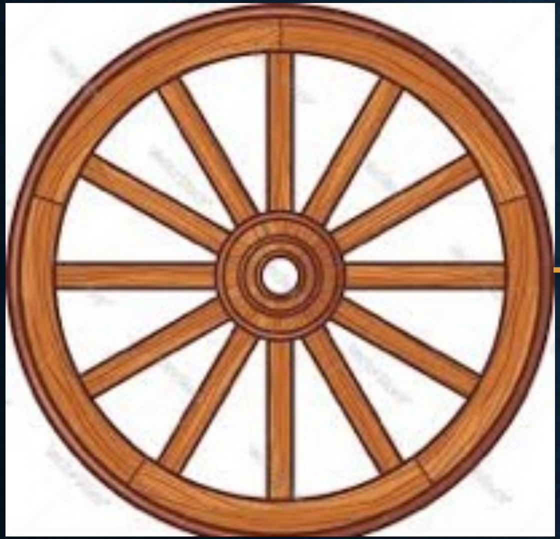
- ***4) Fingerprinting, calibration and Data-collection time issues.***

- ***5) Dynamism – ever expanding system***

- ***6) Translation, orientation And rotation variance***



WHY RE-INVENT THE WHEEL, WHEN YOU CAN UPGRADE IT ?



Existing – Indoor based apps for UCLA

- **Hardcoded routes from one room to another**
- **Missing Real Time Location detection, so basically like *"paper maps but on phone"* ….**

Pros : Nostalgic Cartography Feels
Cons : Stone age Cartography Feels

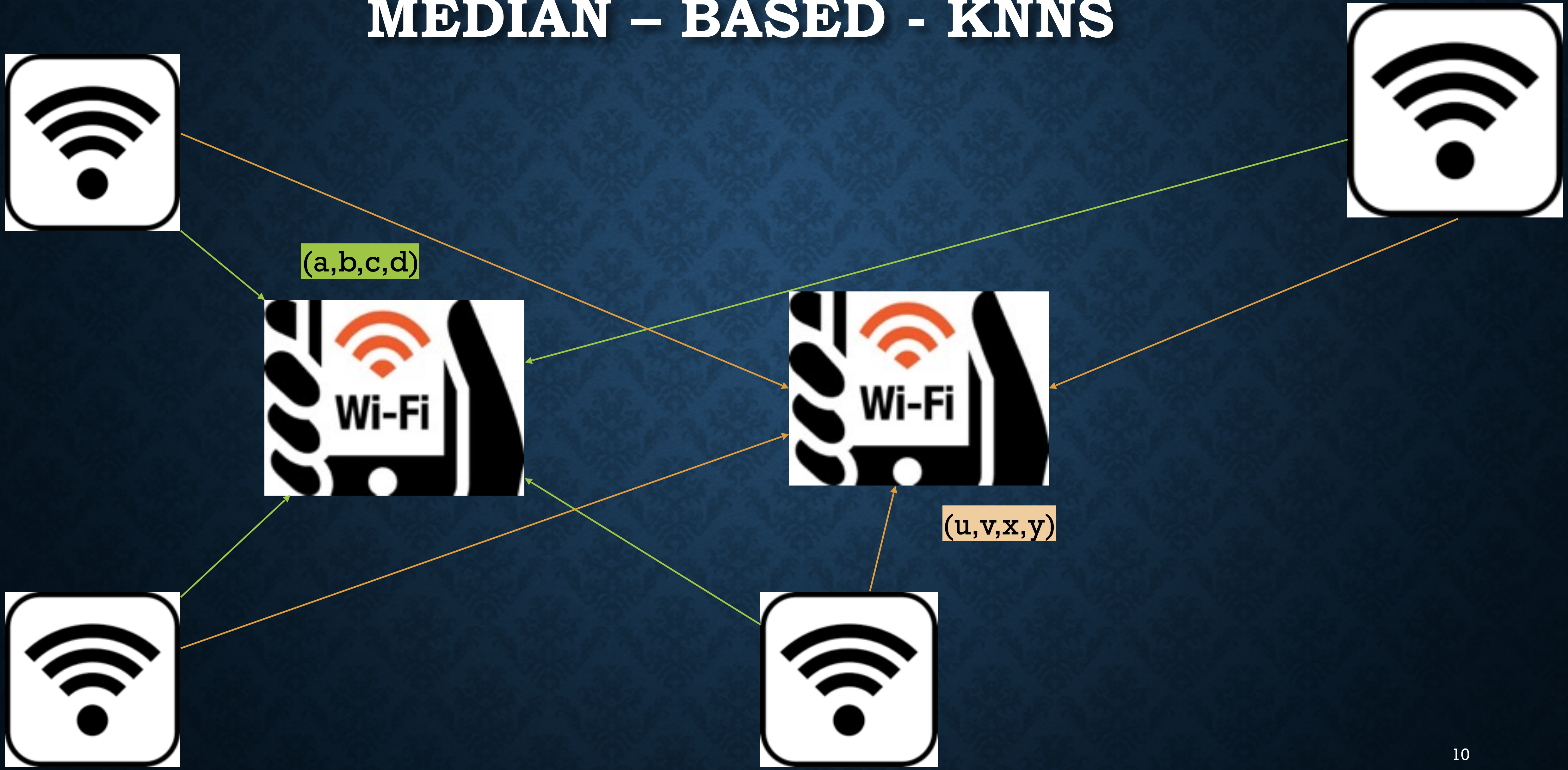
RESEARCH PROJECT VS RESEARCH PRODUCT

- Testing conditions kept as diverse as possible from training conditions.
- Need to account for a wider range of hardware.
- Need to account for future scalability.
- Solving the dilemma of "Need more user data for a better app, but need a better app for more user data"

FREELOC : AN ALGORITHM BY UCLA

- Bases on similarity in peak values of RSS over different time windows.
- Used mean of values within a specified width of peak.
- Tested on different granularities.
- Reduced the test time to 1 min.

MEDIAN – BASED - KNNS



CHALLENGES SCORE-BOARD

- *1) Multiple sources of reflection – so different values of signal strength at the same location.* ✓

  2)Wi-Fi signals degrade with environmental conditions ✗

  3)Wi-Fi Strengths vary between different models of smart-phones ✗

  4) Fingerprinting, calibration and Data-collection time issues. ✗

  5) Dynamism – ever expanding system ✓

  6) Translation, orientation and rotation variance ✗

MEDIAN BASED RANKING KNNS

CHALLENGES SCORE-BOARD

- *1) Multiple sources of reflection – so different values of signal strength at the same location.* ✓ *
  2)Wi-Fi signals degrade with environmental conditions ✓
  3)Wi-Fi Strengths vary between different models of smart-phones ✓
  4) Fingerprinting, calibration and Data-collection time issues. ✗
  5) Dynamism – ever expanding system ✓ *
  6) Translation, orientation and rotation variance ✗

**** Larger block sizes, less granularity and to improve the model over time one has to retrain the entire model***

TOWARDS PERFECTION ….

- **Step 1 : Randomization**

Work with probabilities, The reasoning behind this is that the varying RSS strengths are not ‘random’ as perceived by most literature. Seeing trend of the RSS strengths , we come across the conclusion that the variance comes from average reflections and denoising from various environment and architectural conditions. The architecture factor is more or less fixed (unless a new wall is made, which too is handled later). The environmental factor is solved by ‘Comparision-Techniques’. Overall, out of the many RSS values , random values are chosen multiple times and the class with highest probability is output as the location.

TOWARDS PERFECTION ….

- Step 2: Comparison
- Ranking KNNs had a few advantages but had a loss in granularity because of loss of information, which can be solved by seeing comparision rather than ranks.

MEDIAN BASED RANKING KNNS

Router A

Router B

A>>B

A>B

A=B

A<B

A<<B

TOWARDS PERFECTION ….

- **Step 3 : Optimism**
  - **Here we assume the best possible scenario for every location to be the answer. We tend to be optimistic , and try to find the minimal possible penalty for every location assuming it to be the right answer. The location with the least optimistic penality is chosen.**

TOWARDS PERFECTION ....

- **Combining the ingredients , we propose a self designed Machine Learning Algorithm:**

(and yes, we weren't very creative in naming it) :

- **Optimistic Inverted Comparitive KNNs**

CHALLENGES SCORE-BOARD

- *1) Multiple sources of reflection – so different values of signal strength at the same location.* ✓

  2)Wi-Fi signals degrade with environmental conditions ✓

  3)Wi-Fi Strengths vary between different models of smart-phones ✓

  4) Fingerprinting, calibration and Data-collection time issues. ✓

  5) Dynamism – ever expanding system ✓

  6) Translation, orientation and rotation variance ✓

PESSIMISTIC TESTING FOR OPTIMISTIC ALGORITHM

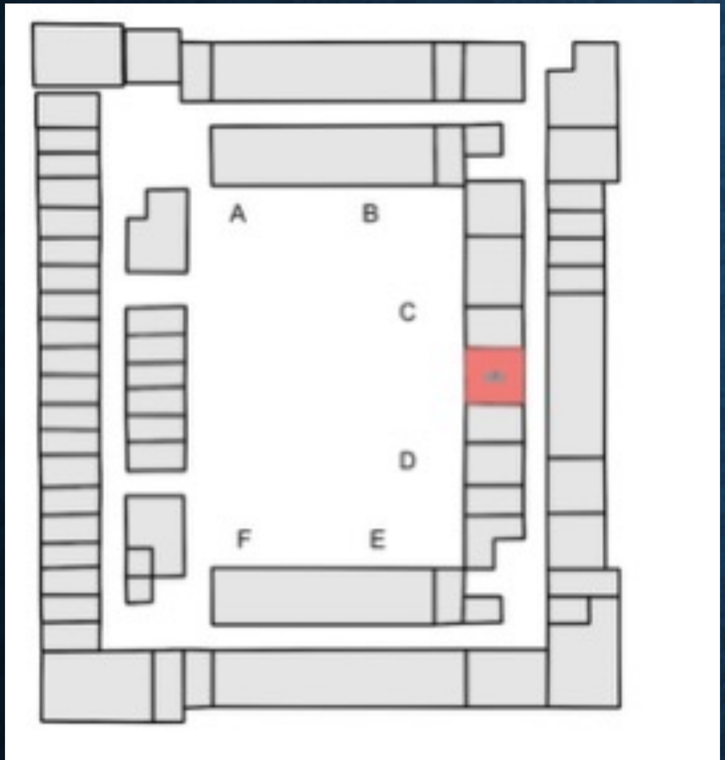
- Everything that could go wrong with localization due to Wi-Fi signals were tried, empty class-rooms, rotating and throwing phones, walking while logging , testing at a different time compared to training. A client (android)- server (php-windows) based android application was made and so you can see it working in this video :

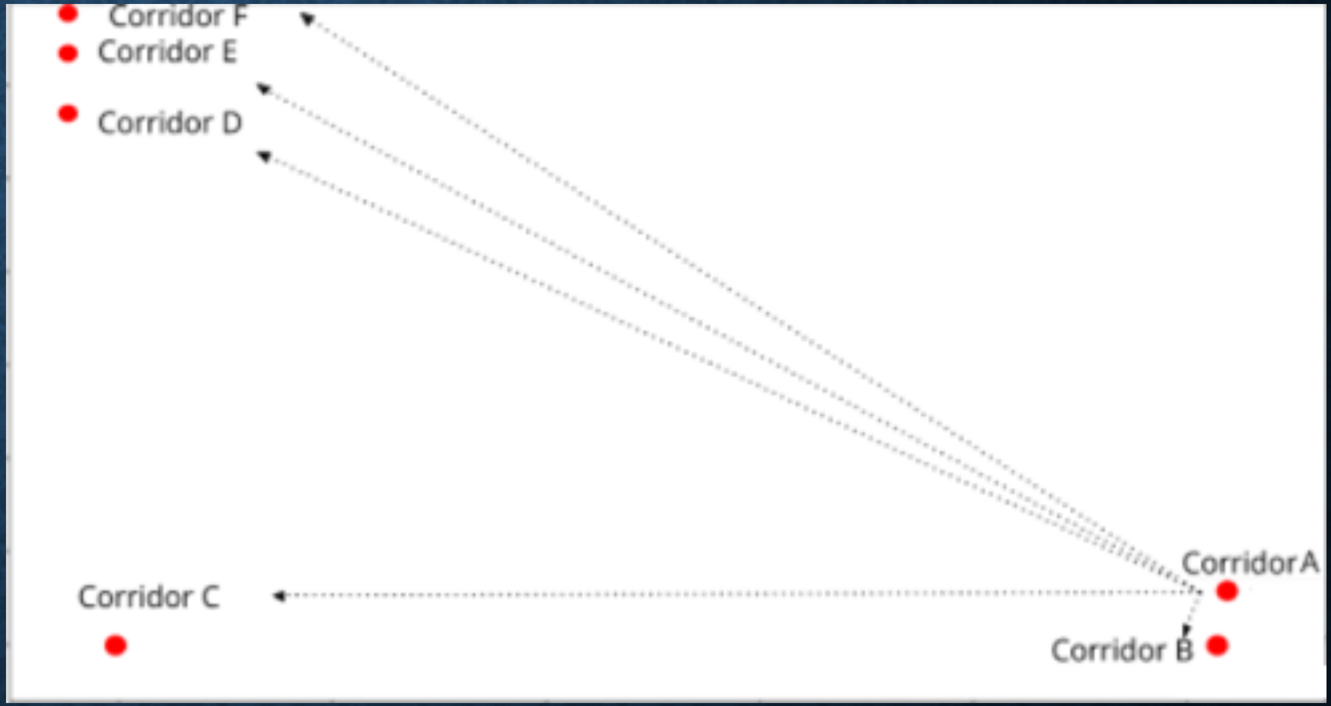
MAKING THE WORK REUSABLE

- Since we have a good approximation of the distance, we tried a method to generate confidence , by predicting the 'neighbourhood' of a point.
- The overlap between neighbourhood defines the confidence we have in our predictions.
- Currently we are using it to improve our performance as by using neighbourhood we have implemented a hierarchical structure of floor-prediction and then room-prediction in the predicted floor-space only.
- In future, the confidence is also important for other clients / developers that may want to hit our service and get a location with a confidence of how much that prediction is true,
- How they use the confidence varies from application to application.

IMPROVING WITH TIME : DYNAMISM AND CROWDSOURCING

- With no additional training time, whenever a new router is introduced , additional training is as simple as adding the new recorded data – and predicting based on the old one. A missing router too does not seem to effect our accuracy at all. This makes our algorithm super scalable.
- ***Visualization of our algorithm shows that it is indeed very effective in predicting distances between class-rooms , hence using clustering , multiple data could be collected at once for a given location – then using re-enforced learning, an update by one phone on the location would result in massive increase in training data.***
- The nature of the algorithm makes it simple to limit the training data size (using hit ratios)


A
B
C
D
F
E


Corridor F
Corridor E
Corridor D
Corridor A
Corridor C
Corridor B

PERFORMANCE IN BOELTER HALL, UCLA

Number of classes : 33

Per-Class Samples : Around 30, 6 seconds samples

Exact Prediction Accuracy : 87.58%

One block Prediction Accuracy : 97.04%

(Training and testing on different phone models, testing done 20 days after training, and distributed throughout the day at different time)

[For comparison, Samsung's paper on Wifi-Localisation published in March 2018 achieves an accuracy of 86% in just 4 locations]

Current System deployed on Boelter Hall - 3rd, 4th, 5th Floor

QUESTIONS AND COMMENTS

Thank you.