ESTIMATING CROWD SIZE USING MOBILE DEVICE ACTIVITY

Jordi Estivill-Dredge, Supervisor: Dr Mark Schulz

PROJECT

Importance

Analytics of human movement is used to identify dwell times, queue times and person counts. This data can highlight how people use a space. With proper analysis, it allows optimisation of a space and its services. For example, airports can identify peak periods, retail can roster more effectively and public transport can improve services.

Current solutions tend to be active or intrusive. Turnstiles, pressure mats or surveys require participation from individuals. Therefore, an approach which doesn't require active user participation is highly desirable.

Description

A number of passive methods have been proposed as a method for estimating crowds. However, people have mobiles, tablets and laptops with them everyday. These devices are frequently checking whether they can connect or use mobile services. Based on this activity, is it possible to estimate crowd size?

BACKGROUND

Technology choice

A variety of technologies have already been investigated to analyse crowds. These include Bluetooth [1], Wi-Fi [2] and cameras [3].

Wi-Fi was chosen for this project because in contrast to competing technologies, it has larger range, works in a variety of environments and has more widespread usage.

SYSTEM DESIGN

A Raspberry Pi installed with Wi-Fi monitoring software Kismet was used to intercept packets. A Python script read the devices' attributes (MAC and RSSI) and transmitted the data. A logging and graphing package received the data.

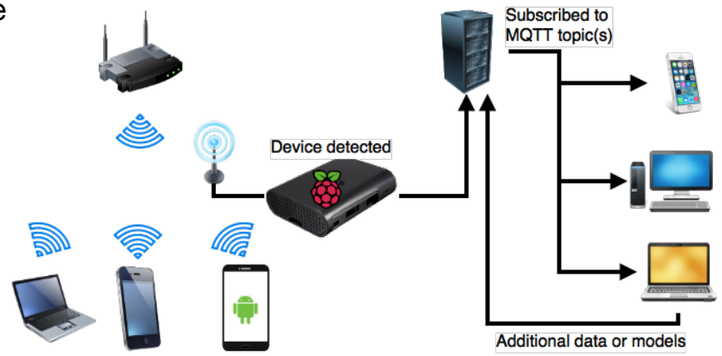
The logs were then merged in Excel and analysed with Weka (a data mining and analysis toolkit).

EXPERIMENTS

A number of trials were undertaken to gather data with different setups and parameters.

All experiments were run in the same location, on varying times and days. The experiments undertaken were:

1. Unfiltered RSSI.
2. Filtered RSSI.
3. Unfiltered RSSI, with manufacturer breakdown.
4. Filtered RSSI, with manufacturer breakdown.



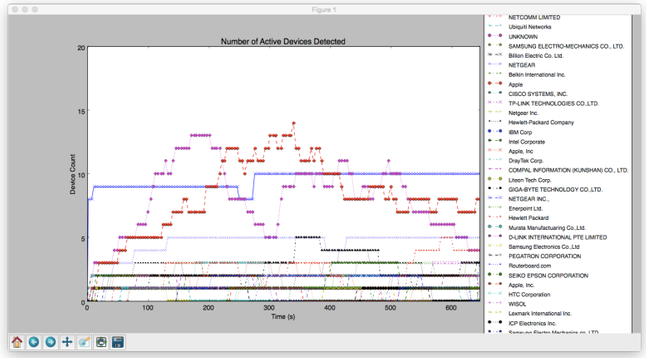
RESULTS

Decision trees were used as the primary analysis tool due to their ability to identify information gain. They also do not assume any relationships between variables. The decision trees constructed in Weka had the following attributes:

1. Unlimited RSSI – 45 nodes, R^2=0.62.
2. Limited RSSI – 53 Nodes, R^2=0.79.
3. Unlimited RSSI, manufacturer breakdown – 275 Nodes, R^2=0.95.
4. Limited RSSI, manufacturer breakdown – 100 Nodes, R^2=0.85.

ANALYSIS

It was initially expected that the 3rd experiment with unlimited RSSI would be the best for estimating crowds.



However in a practical sense, 275 nodes is not easy to follow or implement. Even 100 nodes in the 4th experiment could be considered impractical. Additional analysis tools, or more aggressive pruning algorithms may be required to ensure these decision trees are viable.

The first two experiments, while not as accurate, were certainly more practical.

CONCLUSION

While the latter two experiments were more accurate, they have severe practical limitations. Consequently, other algorithms or tools may be required to improve their viability.

The results found were in line with studies that employed competing technologies. The use of device specific parameters in this study, however, achieved equal to greater accuracy compared to technologies previously researched. (Bluetooth has R^2=0.8 [2]). Therefore, there is reasonable evidence to suggest that mobile device activity can estimate crowd size as predicted.

[1] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," ed, 2013, pp. 193-200.

[2] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. Martin, "Measuring human queues using WiFi signals," pp. 235-238.

[3] U. Stilla, E. Michaelsen, U. Soergal, S. Hinz, and J. Ender, "Airborne monitoring of vehicle activity in urban areas ".