

# Exploratory Navigation for Runners Through Geographic Area Classification with Crowd-Sourced Data

David McGookin

Department of Computer Science  
Aalto University  
Helsinki, Finland  
david.mcgookin@aalto.fi  
www.davidmcgookin.net

Dimitra Gkatzia, Helen Hastie

School of Mathematical and Computer Sciences,  
Heriot-Watt University  
Edinburgh, UK  
{d.gkatzia, h.hastie}@hw.ac.uk  
www.macs.hw.ac.uk/~{dg106, hh117}

## ABSTRACT

Navigation when running is exploratory, characterised by both starting and ending in the same location, and iteratively foraging the environment to find areas with the most suitable running conditions. Runners do not wish to be explicitly directed, or refer to navigation aids that cause them to stop running, such as maps. Such undirected navigation is also common in other ‘on-foot’ scenarios, but how to support it is under-investigated. We contribute a novel method that uses crowd-sourced venue databases to rate a geographical area on its suitability to run in using linear regression. Our regression model is able to accurately predict the suitability of an area to run in (Pearson  $r=0.74$ ) with a low mean error ( $RMSE=1.0$ ). We outline how our method can support runners, and can be applied to other undirected navigation scenarios.

## Author Keywords

Exploratory Navigation; Running; Machine Learning; Regression Analysis; Foursquare; Pedestrian Navigation; OpenStreetMap

## ACM Classification Keywords

H.5.m. Information Interfaces and Presentation: Miscellaneous

## INTRODUCTION

Running is an increasingly popular fitness activity, receiving significant attention from the HCI community [5]. However, how runners navigate the environment, and how this can be supported, has received little attention. This is surprising given recent interest in supporting ‘on-foot’ pedestrian navigation. Such navigation is much more diverse [1] than the turn-by-turn approach adapted from car GPS systems, which are often seen as constraining and dictatorial by pedestrians [6]. This has led to the development of novel techniques that move from explicitly routing users, usually by providing feedback to indicate the bearing and distance towards a known destination [6, 10], or helping them find a longer, but

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.  
*MobileHCI '15*, August 25 - 28, 2015, Copenhagen, Denmark  
Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-3652-9/15/08\$15.00  
DOI: <http://dx.doi.org/10.1145/2785830.2785879>

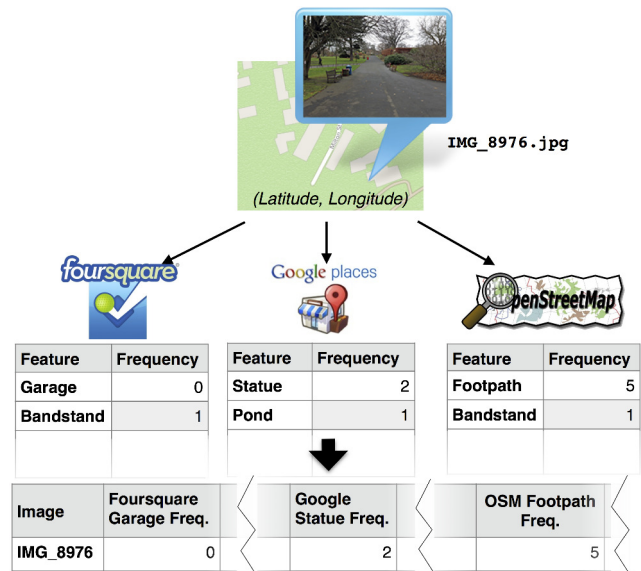


Figure 1. An overview of how the data vector was collected. For each image location, the frequency of venue types within a 150m radius were downloaded. The venue type categories were supplied by each of the services we used (Foursquare, Google Places and OpenStreetMap).

more scenic route [8] to their destination. These approaches have been extended [7, 9] to provide an awareness of multiple points of interest around a user, which can be explicitly or implicitly selected and navigated towards. More recently, work such as Komminos *et al.* [3] has investigated a virtual auditory ‘tether’ to a turn-by-turn route, allowing users to wander from the route but retain awareness, and re-join at a later point.

However, if the user’s objective cannot be expressed as a geographical location, or the user does not wish to define one, or the destination is also the start location, existing A to B solutions fail. For example, a traveller may wish to ‘wander’ around their hotel. They might not have, or wish to have, a clear location to reach, but would like to see interesting things, rather than wandering through an industrial estate. Such A to A navigation scenarios are common in a variety of ‘on-foot’ navigation [1], but supporting them has not yet been investigated.

## KEY CONTRIBUTIONS

In this paper, we focus on how to support this undirected navigation for runners. To our knowledge our work is the first to



Figure 2. Examples of images used in the study. From Left: A pedestrianised area, a shopping street, a park, a residential area and an industrial area.

support this and makes the following key contributions: 1) A novel method that uses machine learning to rate geographical areas according to their suitability for a particular activity. 2) The results of applying our method to running, showing high accuracy and goodness of fit.

### RUNNING NAVIGATION

McGookin & Brewster [4] carried out detailed questionnaires and interviews with runners on how they planned and executed their runs. They found runners often don't pre-plan or carry navigational aids such as maps, even when in new or unfamiliar locations, and do not follow explicit routes. Not doing these being seen as a key motivation for running: freedom. Similarly, pre-planning routes violated the freedom to dynamically change and alter running as the run progresses (e.g. if the runner felt he or she wanted to run farther or shorter distances). The act of running was seen as the primary 'task', with *where* it took place as being largely irrelevant.

However, where runners ran had significant impact on the enjoyment and quality of their run. If the environment contained features requiring runners to compromise their running, for example slowing down to safely cross busy road junctions or navigate through crowds of people, runners sought to find a better environment that would enhance their running experience, such as a park or other quieter areas. Such changes were done in an *ad-hoc* manner whilst running. Running navigation is, therefore, an iterative environmental 'foraging' exercise, with runners constantly seeking to maximise running time in 'good' areas whilst minimising time in 'bad'. When running in familiar locations runners have a good cognitive map of the area, and can make good '*on the fly*' decisions. GPS running watches were commonly used *post run* to contribute to this. *En-route*, these were only used to monitor pace and time. However, when running in unfamiliar places, no such cognitive map is available. Runners will run only if they can immediately see a good location, using this as a starting point to engage in the 'foraging' previously described. If such a location cannot be viewed, many would not run at all.

Such techniques are very different to those used when trying to reach a known destination. Effective solutions must enhance the foraging activity runners exploit, supporting awareness of nearby good and bad areas he or she might be unaware of. However, is it possible to determine the suitability of areas around the user for running, and if so how?

### STUDY DESIGN

To address these issues, we carried out a novel machine learning study to identify a regressor that could be run over a given geographic area to derive a measure of the suitability of that

area for running. Regressors take a data vector that describes relevant conditions of the phenomenon we want to predict (in our case geographic features) and a response variable (in our case a 'suitable to run' score). The regressor is then trained to predict the response variable from the data vector. We can then use this as a model to predict the suitability for running in any area where we can derive a data vector.

### Area Features

To build such a regressor, a data vector must first be derived for each geographical area, and then rate each based on the area's suitability to run in, or not. Based on McGookin & Brewster [4], we identified a total of 84 geographic locations from 3 UK cities in five categories (see Figure 2) that were indicated to influence the suitability of an area to run in. At each location a photograph was taken. We ensured each was representative of its location, and were as similar in style as possible, representing the area from the perspective a runner might encounter it. All were taken over one week, on weekdays only between 11am-5pm. Whilst we could have used Google Street View images (e.g. like [8]), these are taken from a driving perspective (e.g. in the middle of the road in traffic) and do not include many of the (non-drivable) parks or paths indicated as good running locations [4]. When rating the images (see later), we wanted to simulate as closely as possible the view a runner would have.

### Deriving the Data Vector

Given that runners describe good and bad running locations in terms of the types of places and features of the environment [4], we considered that the increasing number of crowd-sourced venue databases may provide suitable data. We could have used data from run logging apps (e.g. runkeeper<sup>1</sup>) which allow runs tagged as public to be download via an API. However, only a small minority, around 10% of runs, are marked as public and accessible<sup>2</sup>. This provides good coverage of popular areas (good places), but allows us to say little of other areas. Are they undesirable for running, or just not used by many runners? We selected Foursquare<sup>3</sup>, Google Places<sup>4</sup> and OpenStreetMap<sup>5</sup>, as these are three of the most popular services and the database of each has a slightly different focus. Foursquare can have venues added by any user, which can lead to duplication or "made up" places [2]. Google focuses on verifying business owners add information - emphasising

<sup>1</sup> [www.runkeeper.com](http://www.runkeeper.com)

<sup>2</sup> <http://tumblr.co/ZTD7ps17-J-ZY>, accessed 5 Feb 2015)

<sup>3</sup> [www.foursquare.com](http://www.foursquare.com)

<sup>4</sup> <https://www.google.com/business/>

<sup>5</sup> [www.openstreetmap.org](http://www.openstreetmap.org)

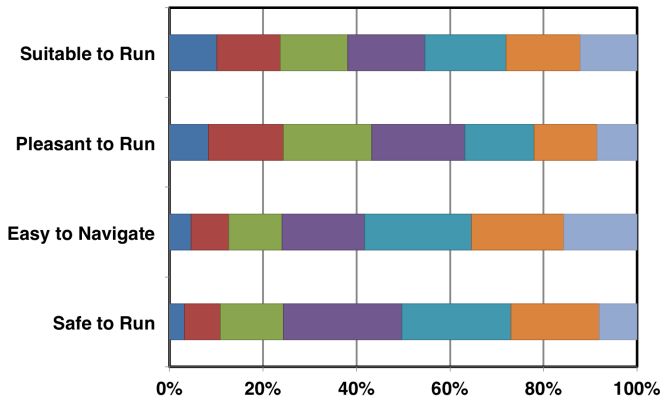


Figure 3. Distribution of Likert ratings for each of the four scales.

commercial premises. Whilst OpenStreetMap focuses on cartographic data, with lower coverage of other venues.

The geo-coordinates of each image, corrected to ensure they matched the taken location of the photograph, were used as query parameters for a search using the public APIs for each of the three services. For each service, we downloaded all **venue types** (for example, Restaurant, Shop, Primary Road, etc.) and their **frequency** within a 150m radius circle centred on the image geo-coordinates. We used the venue types defined by each of the services we used. All three services use a hierarchical model to describe the type of a venue (e.g. a store can be further broken down into a convenience store, clothes store, etc.), and as such we can consider venues at different levels of detail. Both OpenStreetMap and Google Places have two levels of this type classification. For both we decided to use the most detailed level. Thus in OpenStreetMap, primary ‘main’ highways were considered as different venue categories than country tracks, rather than being combined in the higher level ‘highway’ category, as we felt these would be considered differently in terms of traffic volume and thus potential suitability to run in. Foursquare has a much deeper and more detailed hierarchy, and we could not select a single level as with OpenStreetMap and Google Places. The hierarchy was unbalanced, with each level sometimes being too detailed (e.g. ‘Bangladeshi Restaurant’ would be separate from ‘Bengali Restaurant’) for our goals, or were too general (e.g. ‘school’ would combine ‘flight school’, ‘circus school’ and ‘elementary school’). As such, each venue instance was added to each of the foursquare venue categories it appeared in (e.g. adding an elementary school to both the ‘school’ and ‘elementary school’ category). The three data vectors derived from each service (Foursquare, Google Places and OpenStreetMap) were combined into a single data vector for each of the 84 images (see Figure 1).

### Subjective Ratings of Suitability

To train the regressors, each of the data vectors (and their represented areas) must be manually rated as to their suitability for running. To be meaningful, this must be done by runners. We generated an on-line questionnaire on SurveyMonkey<sup>6</sup>. Participants provided brief details of their running (runs per week, distance and whether they ran with others or alone),

<sup>6</sup>www.surveymonkey.com

Regressor	$r$	$R^2$	RMSE
MLR	0.714	0.50	1.30
M5P	0.709	0.50	1.31
M5Rules	0.710	0.50	1.03
M5Rules + personal	0.740	0.54	1.27

Table 1. Regression results for predicting ‘suitable to run’ using Multi-variable Linear Regression (MLR) and the two decision tree (M5P and M5Rules) regressors. Also shown is the result of M5Rules when including personal running information.

before rating a subset of 28 images (out of our total set of 84) in a random order. Based on the findings of McGookin & Brewster [4], participants rated each image on 4, 7-point likert scales: how suitable the area was to run in, as well as its pleasantness, ease of navigation and safety. Participants were instructed to base their rating solely on the image shown. Such image rating would, therefore, be similar to the quick visual judgement runners would have to make when running. The questionnaire was distributed to running clubs and on-line running forums based in the U.K. Completed responses were entered into a draw for a £50 running shop voucher.

### DATA COLLECTION AND REGRESSION ANALYSIS

242 runners (139 male) completed the questionnaire, leading to 5,327 image ratings. Figure 3 shows the distribution of ratings over all images for each of the four attributes. High consistency was found between the ratings for individual images (Chronbach  $\alpha = 0.9$ ). Therefore, we can consider the images, questions and scales used to rate them are appropriate.

To determine if we could predict the ‘suitability to run’ response variable from the vector, we applied three regressors from the Weka machine learning toolkit<sup>7</sup>: Multivariable Linear Regression (MLR) and two decision trees (M5P and M5Rules) [11]. Each was run with 10-fold cross validation. Here the 5,327 responses are randomly split into 10 groups. The regressor repeatedly builds a model using 9 groups as training data, evaluating this by predicting the response variable for the 10th group. This means the regressor is not tested on the same data it was trained on. The goodness of fit of the model to predict the response variable is calculated as an average across all iterations, and taken as the overall accuracy of the regressor to predict the response variable. Such validation is a ‘gold standard’ in Machine Learning [11].

All three regressors produced accurate models for ‘suitable to run’, with no significant variation between them (see Table 1). All provided strong Pearson correlation coefficients ( $r$ ) between the regressor’s predicted score and the rating provided by runners. The Root Mean Square Error (RMSE) showed the mean error between subjective and predicted values was on average 1 point on our 7-point Likert rating (e.g. an area with a ‘suitability to run’ of 1 rated by runners, predicted as 2 based on the model). Runners are largely concerned with the binary difference between ‘good’ and ‘bad’ areas [4], so an error of 1 Likert point is adequate to support this.

Inspecting the generated M5Rules (this has a lower *RMSE*) model allows one to see which features are indicative of good and bad areas. Areas with large concentrations of shops -

<sup>7</sup>http://www.cs.waikato.ac.nz/ml/weka/

indicative of busy retail areas - were strongly negatively correlated with the ‘suitable to run’ response variable. As were establishments indicative of industrial areas, such as car dealerships. Outdoor establishments - such as universities, and art galleries with outdoor space, parks, as well as cycleways, footpaths and tertiary (minor) roads - had a strong positive contribution. Residential areas - with few dedicated pedestrian crossings - were also rated highly.

### Personalising the Model

The Coefficient of Determination ( $R^2$ ) shows the regressors can account for around 50% of the variation between the response variable and its predicted value. Much of this is likely due to individual differences between raters. This indicates we can “build out” at least some of the variation by tailoring the model to each user. To test this, we re-ran the m5Rules regressor including the responses on running habits each runner provided when rating the images (gender, times running per week, distance ran and whether they ran with others or alone) as part of the data vector. Including these showed a slight increase in the correlation coefficient (see Table 1). Although this is a *post-hoc* consideration, and our running habits questions were not intended to be used in this way, it indicates personalisation as a fruitful avenue for future investigation.

### Inferring Suitability to Run

To apply our method, images must be manually rated by domain experts. In this case runners. However, it may be difficult to find enough experts to do this, and in any case rating is time consuming (our image ratings took two months to complete). As well as rating the ‘suitability to run’ of an area, participants also rated each image on three other attributes (pleasantness, navigability and safety) discussed as contributing to suitability [4]. These are simpler, and potentially do not require domain experts. These sub-ratings had a medium to strongly positive correlation with each other and ‘suitability to run’ (Spearman’s Rho between 0.52 and 0.81,  $p < 0.05$ ). To determine if we could predict ‘suitability to run’ from them, we applied the M5Rules regressor again, but used the three sub-ratings as the data vector with ‘suitability to run’ as the response variable. Validated with 10-fold cross validation, an accurate model with good fit (Pearson  $r=0.86$ ) and low error ( $RMSE=0.99$  on our 7-point Likert scale) was determined.

This raises the possibility that we can define a weighting function over a set of “atomic” navigability characteristics. Images could be rated without the need for domain knowledge (e.g. via Amazon Mechanical Turk), with these used to build the regressor models. Domain experts would then only need to provide the weighting of these “atomic” attributes. This would require us to uncover the full set of “atomic” navigability attributes. The three we used work for running, but there are likely others more relevant to other undirected navigation scenarios. However, it would allow models to be built from much larger data sets and significantly cut down the amount of work needed to apply our method to new scenarios.

### Service Influence

In applying machine learning, there is a real world cost associated with collecting each attribute of the data vector (e.g.

Service	$r$	$R^2$	RMSE
OSM	0.709	0.50	1.31
FQ	0.702	0.50	1.33
GOOG	0.709	0.50	1.31
ALL	0.710	0.50	1.03

**Table 2. M5Rules results based on groups of features Foursquare (FQ), Google Places (GOOG) and Open Street Map (OSM).**

each value may derive from a different sensor that consumes power on a mobile device). In our case the cost is not in the number of attributes, but rather the number of services they are derived from. Each of the three services we used requires a separate network request. Each request has a greater overhead than the few bytes we download. As we do not know where users will run, we cannot pre-cache data, and we also need to download new data as the runner runs. Therefore minimising this overhead is desirable.

To determine the individual impact of services, we split the combined data vector back into a Foursquare, Google and OpenStreetMap part and re-ran the regressors on each. Table 2 shows results for the M5Rules regressor on these split vectors. Interestingly, this resulted in no reduction in accuracy, with the regressors producing models with the same accuracy and goodness of fit using individual services as they could with the combined vector of all three. This indicates sufficient level of overlap in the coverage of features across these services to accurately model the area. Adding additional services (e.g. Yelp ([www.yelp.com](http://www.yelp.com))) is likely to add further cost, without improving performance. However, we could also use this to increase the robustness of a real world system. If one service goes down we can switch to a model derived from another.

### CURRENT LIMITATIONS

Our work shows that venue databases can be used to predict the suitability of an area for a particular activity (running). We chose images across three different cities in the U.K., and our regressor was able to determine the scores of images from one city using the results of another to high accuracy. This gives confidence that our results are generalisable to at least U.K. cities. However, we do not currently know if the geographical predictors that were identified as important in the U.K. will also apply to different countries.

The suitability of an area to run through may also change over the day. We took each photograph at around the same time of day on a weekday. However, a pedestrian shopping area may be a bad place to run on a busy shopping day, but become a good area once the shops are closed. Similarly, a quiet area during the day may become an area where users feel less safe after dark. How to incorporate these temporal aspects into the model is an interesting research challenge. It is, for example, unrealistic to take images for all times and days, as there are not enough users available to rate these images.

### FUTURE WORK

Such issues, although important, are beyond the scope of our study here. To answer them, we are planning long-term field studies with runners, using smart-watches presenting novel visualisations of the area around the user (see Figure 4),



Figure 4. A potential visualisation to present the running suitability of the nearby environment to users, using the regressor to generate segmented pie charts that egocentrically show the suitability of good (green), neutral (yellow) and bad (red) running areas around the user. The blue segment indicates the direction back to where the user started running.

driven by our model, that allow quick glanceable understanding to support foraging behaviour. For example, thresholding the 7-point area score in three bands (good, bad and neutral) to present a heat map of suitable running locations around the runner's current location. Such visualisations will provide the runner with glanceable information that supports, rather than tries to replace the exploratory nature of running. Through both data logging and audio memos (generated by tapping the watch screen), we will be able to collect data to determine how both our model applies to different countries and to temporal aspects. Through doing so, we will be able to tailor our model to better fit and predict running suitability.

## CONCLUSIONS

Our results show it is possible to build an accurate regression model, using existing pervasive crowd-sourced media, to predict the suitability of a geographical area to support a particular activity (running). Our model validated with a Pearson coefficient of 0.74, and with a low RMSE of 1 on the 7-point Likert rating scales we used. This model is therefore more than robust enough to determine the binary (good/bad) determination of an area that runners make when running [4].

Our work is the first to consider how we might support navigation through the environment when the user does not have, or wish to have, a defined destination, and therefore cannot be routed towards it. Although we have looked specifically at running, there are a significant number of similar navigation scenarios where it is the journey that is more important than a destination. The novel method that we have used here could be employed to support these activities as well. We hope our work will support and encourage further investigation into supporting the full range of 'on-foot' exploratory navigation humans engage in.

## ACKNOWLEDGMENTS

We would like to thank all of the runners for participating in the image rating study. This work was in part supported by the EPSRC GUI project (EP/L026775/1).

## REFERENCES

1. Allen, G. L. Spatial abilities, cognitive maps, and wayfinding. *Wayfinding behavior: Cognitive mapping and other spatial processes* (1999), 46–80.
2. Cramer, H., Rost, M., and Holmquist, L. E. Performing a check-in: emerging practices, norms and 'conflicts' in location-sharing using foursquare. In *Proc. of CHI 2011*, ACM (Vancouver, Canada, 2011), 57–66.
3. Komninos, A., Barrie, P., Stefanis, V., and Plessas, A. Urban exploration using audio scents. In *Proc. of MobileHCI 2012*, ACM (San Francisco, CA, USA, 2012), 349–358.
4. McGookin, D. K., and Brewster, S. A. Investigating and Supporting Undirected Navigation for Runners. In *Ext. Proc. of CHI 2013*, ACM (New York, NY, USA, 2013), 1395–1400.
5. O'Brien, S., and Mueller, F. Jogging the Distance. In *Proc. of CHI 2007*, ACM Press (San Jose, CA, USA, 2007), 523–526.
6. Pielot, M., Poppinga, B., and Boll, S. PocketNavigator: vibro-tactile waypoint navigation for everyday mobile devices. In *Proc. of MobileHCI 2010*, ACM, ACM Press (Lisbon, Portugal, 2010), 423–426.
7. Pielot, M., Poppinga, B., Heuten, W., and Boll, S. Tacticycle: supporting exploratory bicycle trips. In *Proc. of CHI 2012*, ACM (New York, NY, USA, 2012), 369–378.
8. Quercia, D., Schifanella, R., and Aiello, L. M. The Shortest Path to Happiness: Recommending Beautiful, Quiet, and Happy Routes in the City. In *Proc. of Hypertext and Social Media 2014*, ACM (New York, NY, USA, 2014), 116–125.
9. Stahl, C. The roaring navigator: a group guide for the zoo with shared auditory landmark display. In *Proc. of MobileHCI 2007*, ACM (New York, NY, USA, 2007), 383–386.
10. Williamson, J., Robinson, S., Stewart, C., Murray-Smith, R., Jones, M., and Brewster, S. Social gravity: a virtual elastic tether for casual, privacy-preserving pedestrian rendezvous. In *Proc. of CHI 2010*, ACM (New York, NY, USA, 2010), 1485–1494.
11. Witten, I., Frank, E., and Hall, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Morgan Kaufmann Publishers, 2011.