Discovering User Location Semantics using Mobile Notification Handling Behaviour

Andreas Komninos[1,2][0000-0001-8331-1645], Ioulia Simou[2], Elton Frengkou[2], and John Garofalakis[1,2]

[1] University of Patras, Rio 26504, Greece
[2] Computer Technology Institute & Press "Diophantos", Rio 26504, Greece
akomninos, simo, frengkou, garofala@ceid.upatras.gr

Abstract. We analyse data from a longitudinal study of 44 participants, including notification handling, device state and location information. We demonstrate that it is possible to semantically label a user's location based on their notification handling behaviour, even when location coordinates are obfuscated so as not to precisely match known venue locations. Privacy-preserving semantic labelling of a user's location can be useful for the contextually-relevant handling of interruptions and service delivery on mobile devices.

Keywords: Interruption management · Mobile notifications · Semantic location labelling.

1 Introduction

As users of mobile devices roam through urban environments, a wealth of data can be collected from their devices about their current whereabouts and activities. While it is relatively easy to obtain the location of a user, within a given accuracy estimate (e.g. through GPS, connection to Wi-Fi or 4G networks), a harder task is to assign semantics to the user's location. The typical method of resolving this, is by comparing the user's coordinates against a database of known locations, and there are several commercial services that offer this type of information (e.g. Google Places API). Therefore, given a user's location coordinates, it is relatively easy to obtain the venue that a user might currently be at, and therefore to infer their current activity (e.g., they are at Cinema X, and thus quite likely watching a movie). This knowledge is valuable for the purpose of offering contextually relevant notification handling to users. Currently, users are left on their own in terms of how they might manage notifications under different contexts [3]. However, automatic notification management can offer opportunities for a better and more socially aware mobile use experience [2]. Taking the cinema example, a device could automatically suppress incoming notifications which are not relevant at the current location (e.g. [11]) or automatically set the device ringer mode to silent for the duration of the user's stay at that location.

There are several confounding factors to being able to achieve this goal. First, user location coordinates might not be available, or accurate enough to provide a reasonable estimate of venue (e.g. the user might be indoors, or the user might be connected to a sparse 4G network only). Further, the user might be mobile and therefore rapidly moving across venues, hence a continuous lookup of the user's location is required, expending device power and network bandwidth. Even more, for services such as this to work, the user's location needs to be sent to a remote server, potentially compromising user privacy.

As discussed in existing literature, users receive a significant volume of notifications during the day, from on-device events (e.g. network availability, battery status) and external services (e.g. instant messaging), which can reach several hundreds ([12],). These events can become opportune moments for assessing the user's location. The user behaviour in handling these notification events can vary significantly across time (e.g. [8]), and we can assume that the behavioural choices are influenced by the location context and semantics as well, even though there is no previous literature to investigate this. For example, while watching a movie at the cinema, the user might take longer to notice an incoming notification since their device will probably be set to "silent mode" and tucked away, or even if they do, they might chose to ignore it until the show is over.

In this paper, we explore the use of notification handling behaviour and device state information, as an additional source of information for overcoming problems with user coordinate availability and accuracy. Using supervised machine learning algorithms on a dataset of notification and location samples from several users, we predict user location semantics and demonstrate that notification handling behaviour can overcome the problem of location accuracy.

2 Related work

Discovering location semantics is the research effort directed towards assigning categorical labels (e.g. "Home", "School", "Shop") to venues represented in a dataset with at least a set of coordinates (latitude, longitude) and optionally a given name (e.g. "Mike's cafe"). Location semantics are important for a range of location based services, such as point-of-interest (POI) search and recommendation. Commercial applications such as Google Maps, Foursquare and Tripadvisor maintain large databases of POIs, relying largely on users adding and/or modifying these. One issue with this approach is that represented venues are not always correctly semantically labelled by the users, and also the reliance on user effort means that many real-world POIs may be often left out of the service. Previous research has frequently focused on the automatic semantic labelling of locations, with a variety of means. For example, in [6], check-ins from social networks (Twitter) were used to identify users' home locations, with good accuracy. In [9], data from location diary studies was used to build a model to automatically infer user home and office locations, using GPS traces as an input, resulting in reasonable performance for both categories (100% and 66%). In [5], check-in data (e.g. number of visitors, diurnal distribution, stay time etc.) was again used to predict venue labels across 8 categories, however with mixed results across the different categories (F-score betwen 54% - 92%). In [7], a spatiotemporal topic model was used to leverage location "tags" left by users, in order to determine the location category, with an average accuracy $\approx 60\%$.

Other studies have leveraged sensor data in addition to other contextual information for semantic place labelling. In [15], GPS, accelerometer, Bluetooth and Wi-Fi data were used amongst others to achieve an accuracy $\approx 75\%$ across 10 different location categories. Similar results are obtained in [13], where labelled and unlabelled data are used to implement a semi-supervised learning approach to predict across 9 categories. In [4], 11 categories are predicted from, using features related to stay time, device battery, applications used, user current activity etc. Results show an accuracy of $\approx 55\%$ with a range of classifiers.

There are some common themes in the previous literature, which can be identified. First, where multiple classifiers have been used (e.g. decision trees, SVMs, random forests), the results do not seem to vary significantly. Most often, it is the type and number of features introduced to the model which have the most impact. Secondly, a larger number of categories makes the likelihood of misclassifications higher. Both in [4] and in other work such as [10], it is demonstrated that a less fine-grained categorisation approach improves results significantly (e.g from $\approx 65\%$ with 10 categories, to $\approx 89\%$ when these are collapsed into 3). Another issue is that, as demonstrated in most papers (e.g. [4]), there is a significant class imbalance in the datasets used. This is somewhat problematic since in all reviewed works apart from [5], the measure of accuracy is used, which is heavily influenced by the prevalence of certain categories [1]. Hence, comparisons with the performance of these previous approaches is done with some hesitation.

To the best of our knowledge, the use of notification handling behaviour as a feature for semantic place labelling has not been investigated in the past. Hence the goal of our paper is to explore how this information can be used for the task of semantic place labelling.

3 Study methodology

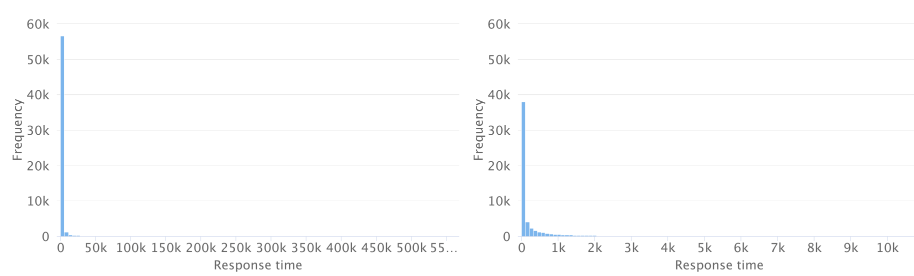
3.1 Apparatus and participants

We developed a UI-less notification logging application for Android devices, which runs unobtrusively on the device as a background service. Using the Android NotificationListener service, which allows the capture of issued notification details, as well as various other Android APIs (e.g. PowerManager, DisplayManager), we collected features about the notifications and the user's device state at the time of issue. We also exploited the Google Places API to retrieve details about the user's presumed location at the time of notification issue. This API requests the user's location coordinates, and returns a list of likely places where the user is located, along with a confidence level. We logged the place which had the highest confidence value. The data features collected are discussed in detail in section 3.3. All data was uploaded to a remote server at frequent intervals during the day, provided the user had wi-fi connectivity.

A call for participation was issued to undergraduate students at our local university. The application was installed on their device, a consent form was signed and participants were instructed that they could quit the study at any time. The study automatically ended after 3 months of use. They were requested to leave location services enabled on their device for the duration of the study, although we did not enforce this condition. In total, 44 participants took part in the study (26 female). From this set of participants, we excluded several participants who participated for fewer than 10 days and who provided fewer than 50 notification log entries, resulting in a subset of 31 participants. Participants provided data that spanned an average of 30.87 days (sd=16.15, min=13, max=84).

3.2 Dataset preparation

In total we collected 204,074 notifications from the users. In the dataset, we noticed that a significant number of notifications (38,400) were issued by the system and immediately dismissed. This phenomenon was observed for all users, although for some users the proportion of such notifications was unusually large. We are not certain why this happens. Further investigation of the package name showed that some system applications might be issuing such notifications (perhaps as a means of interprocess communication), although it might be the case that a user is also manually quickly dismissing some notifications (within the resolution of 1 second). We decided to exclude such notifications from the dataset. Further, we removed from the dataset all notifications for which the "flag" feature values indicated that they were ongoing events and not user-dismissable (e.g. an ongoing phonecall or download). These notifications are automatically dismissed by the system and hence offer no value to our research goal. From the remaining notifications, a significant number did not contain location information, since the user's location services might have been switched off at the time, or the service might not have been available. We also excluded these from the dataset. After these exclusions, the dataset contained 59,221 user-dismissed notifications with location details.



(a) Entire dataset (b) Response time < 3h

Fig. 1: Distribution of response time to notifications (100 bins)

Examining the pruned dataset, we observed that the average response time to notifications is 1,366.93s (sd=11,255.82), with a maximum response time of

562,302s. A histogram of response time to notifications shows a power-law distribution (Fig. 1). Based on this observation, we limited the dataset to only notifications that were attended to within 3 hours of issue, resulting in 57,737 notifications (97.5% of the original dataset). As can be seen, even after culling the dataset further, the distribution of response times to notifications maintains a power-law shape. This finding is consistent with previous works such as [8].

3.3 Dataset features

To address the problem at hand, we used raw and synthetic features obtained from the user's device. To begin, the raw data features collected from users are shown in Table 1.

From these raw features we synthesized a further set of features, to create the final dataset to be used for prediction, as shown in Table 2. Notably, we used the current device ringer mode and programmed notification modalities (custom or default) to determine the true modalities used to deliver the notification, as per [8]. Further, a place can belong to multiple categories. These are reported in a non-ordered list by Google, ostensibly therefore the order of appearance shows the prevalence of a category type (e.g. "Bar, Restaurant, Cafe" shows that a place is primarily of type "Bar", but also functions as a restaurant and cafe). We therefore extract the primary category of a venue. In doing so, we observed that many places included the vague category "Point of Interest". Hence, where this was the primary category, it was replaced by the immediately subsequent category type.

Another note here relates to Google's list of categories, where 127 different categories are listed. Predicting on 127 category classes is possible, but presents an unnecessary complexity to the problem, as many venue categories are quite similar in nature and it can be expected that a user will exhibit similar behavioural patterns in these. For example, "Church" and "Mosque" are both places of worship, where devices are typically kept on silent, and users do not readily engage in notification handling. We therefore attempted to group the individual categories into larger sets, as per Table 3. Ultimately, we assigned to each place the super-category to which it belongs, based on its primary category type. An exception to this were the "Miscellaneous" and "Entertainment areas" categories, since for these the user behaviour might be quite different depending on conditions (e.g. a user probably can't notice a notification in a night club as easily as in a cafe), hence for these we used the primary categories ungrouped. As a result, we find that the user notifications were issued at 24 distinct place categories and distributed unevenly (Table 3, non-grouped primary categories capitalised). Finally, it's important to note that the location coordinates collected by our app, are not the user's actual coordinates, but the coordinates of the venue that is the user's most likely current place, as reported back by Google's API. We do not store the user's actual location coordinates for privacy reasons.

As can be seen in Fig. 2, users receive a varying amount of notifications throughout the day. The distribution is similar to that reported in previous

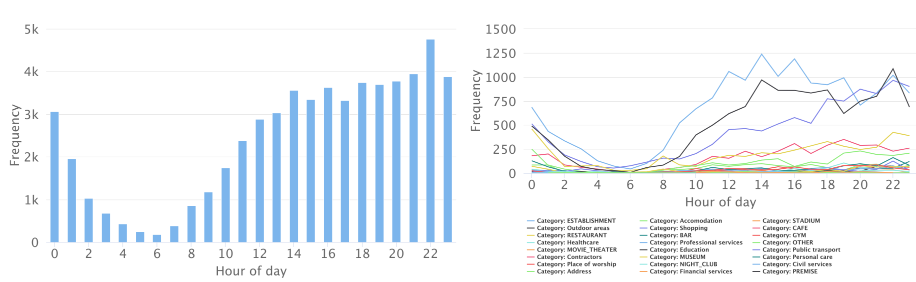
Table 1: Raw data features collected

Notification Details	
Time posted	Unix timestamp of notification issue
Time dismissed	Unix timestamp of notification dismissal
Package name	Application that created the notification
Sound	Whether the notification was programmed to issue a custom sound alert
LED	Whether the notification was programmed to issue a custom LED blink pattern
Vibration	Whether the notification was programmed to issue a custom vibration pattern
DefaultSound	Whether the notification was programmed to use the default sound alert
DefaultLED	Whether the notification was programmed to use the default LED blink pattern
DefaultVibration	Whether the notification was programmed to use the default vibration pattern
Priority	The notification priority category
Notification flags	Additional information about the notification
Device state	
Ringer mode	The current device ringer mode (silent, vibrate only, full)
Idle state	Whether the device is in an idle state
Interactive state	Whether the device is in a state ready to interact with the user (screen on, processor awake)
Lockscreen notifications allowed	Whether notifications are visible from the user's lock screen
Location Details	
Place name	Name of the most likely current place
Place categories	The categories assigned to the most likely current place
Confidence	Confidence of reporting the most likely current place
Latitude	Decimal coordinates of the most likely current place
Longitude	Decimal coordinates of the most likely current place

Table 2: Final feature set

Notification Details		
Response time	Time dismissed - time posted	Synthetic
Hour issued	Hour of day at notification issue [0-23]	Synthetic
Day of week issued	Day of week at notification issue [1-7]	Synthetic
Had Sound	Whether the notification was issued with a sound	Synthetic
Had LED	Whether the notification was issued with a LED blinking pattern	Synthetic
Had Vibration	Whether the notification was issued with a vibration pattern	Synthetic
Priority	The notification priority category	Raw
Device state		
Idle state	Whether the device is in an idle state	Raw
Interactive state	Whether the device is in a state ready to interact with the user (screen on, processor awake)	Raw
Lockscreen notifications allowed	Whether notifications are visible from the user's lock screen	Raw
Location Details		
Place category	The primary place category	Synthetic
Latitude	Decimal coordinates of the most likely current place	Raw
Longitude	Decimal coordinates of the most likely current place	Raw

literature, such as [4]. More importantly, we note that the diurnal distribution varies pronouncedly for only a few categories, whereas for other categories, it remains rather consistent. This is an expected result, since different venue types exhibit different diurnal visitation patterns [5]. Further, we note the distribution of response times to various notifications on a hourly basis (Fig. 3a). The pattern is similar to the findings in [8], showing the distinct user behaviour in handling notifications throughout the day. Distinct response time averages are also noted across the categories (Fig. 3b), demonstrating that attentiveness to the device is likely related to device ringer mode and current user activity (Fig. 4).

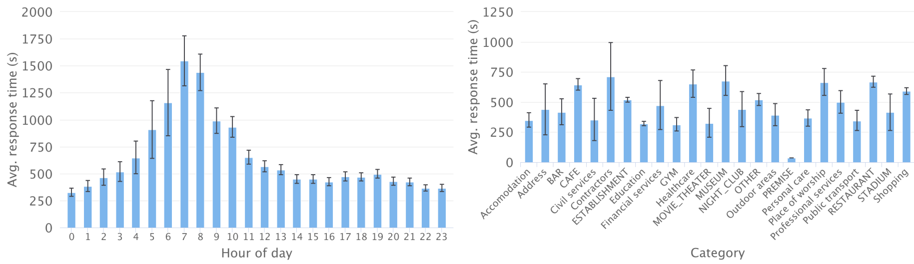


(a) All notifications (b) Per place category

Fig. 2: Diurnal distribution of notifications

Table 3: Grouped place categories

Category group	Categories	Samples
Accomodation	Campground, Lodging, Room, Rv Park	1,350
Address	Administrative Area Level 1, Administrative Area Level 2, Administrative Area Level 3, Country, Geocode, Locality, Political, Post Box, Postal Code, Postal Code Prefix, Postal Town, Street Address, Sublocality, Sublocality Level 1, Sublocality Level 2, Sublocality Level 3, Sublocality Level 4, Sublocality Level 5, Synthetic Geocode	86
Civil Services	City Hall, Courthouse, Embassy, Fire Station, Local Government Office, Police, Post Office	89
Contractors	Electrician, General Contractor, Moving Company, Painter, Plumber, Roofing Contractor	76
Education	Library, School, University	11,996
Entertainment Areas	Amusement Park, Aquarium, Bar, Bowling Alley, Cafe, Casino, Gym, Movie Theater, Museum, Night Club, Restaurant, Stadium, Zoo	11,157
Financial Services	Bank, Atm, Finance	93
Healthcare	Dentist, Doctor, Health, Hospital, Physiotherapist	617
Miscellaneous	Establishment, Floor, Other, Point Of Interest, Premise, Subpremise	18,347
Outdoor Areas	Colloquial Area, Natural Feature, Neighborhood, Park, Parking, Route	516
Personal Care	Beauty Salon, Hair Care, Spa	1,104
Place Of Worship	Cemetery, Church, Hindu Temple, Mosque, Place Of Worship, Synagogue	758
Professional Services	Lawyer, Accounting, Car Dealer, Car Rental, Car Repair, Car Wash, Funeral Home, Insurance Agency, Laundry, Locksmith, Real Estate Agency, Storage, Travel Agency, Veterinary Care	659
Public Transport	Airport, Bus Station, Intersection, Subway Station, Taxi Stand, Train Station, Transit Station	580
Shopping	Art Gallery, Bakery, Bicycle Store, Book Store, Clothing Store, Convenience Store, Department Store, Electronics Store, Florist, Food, Furniture Store, Gas Station, Grocery Or Supermarket, Hardware Store, Home Goods Store, Jewelry Store, Liquor Store, Meal Delivery, Meal Takeaway, Movie Rental, Pet Store, Pharmacy, Shoe Store, Shopping Mall, Store	10,309



(a) Diurnal, all notifications (b) Per place category

Fig. 3: Distribution of response time to notifications

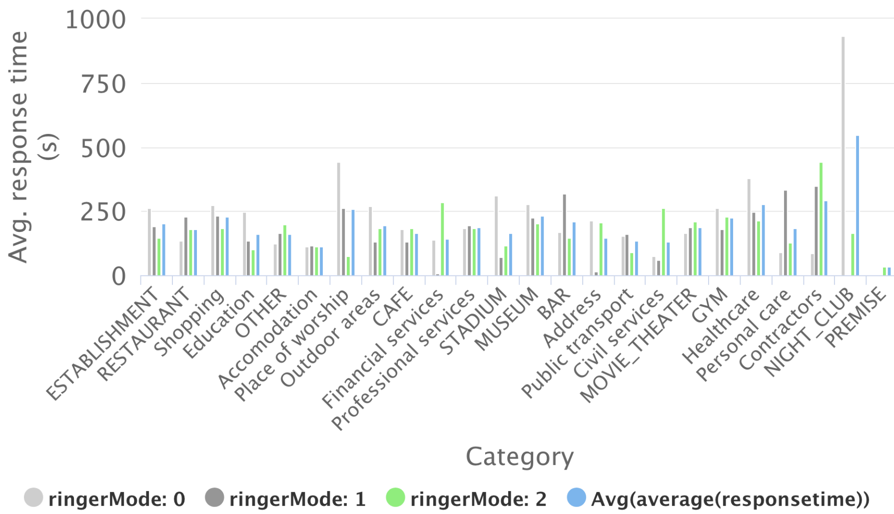


Fig. 4: Average response time per category and device ringer mode (0=silent, 1=vibrate only, 2=all modalities)

4 Predicting user location types

4.1 Algorithms and parameter selection

For our analysis, we used decision trees to perform multinomial classification on the prediction target (place type), since they have been shown to demonstrate comparable performance to other methods [5]. To obtain an estimate of good parameters to use, we employed an evolutionary algorithm search process on a small hold-out dataset. The final parameters used for the algorithm are Maximal depth:23, Minimal gain:0.013, Minimal leaf size:2, Minimal split size:4. Throughout the analysis reported in the following sections, we used a 10-fold cross-validation approach. As per other studies, we note that there is an imbalance in the frequency of location categories (Fig. **??**), hence for performance we adopt the F1-score (macro-averaged), which is more appropriate for imbalanced datasets, compared to the accuracy measure usually encountered in previous literature.

4.2 Decision tree modelling performance

As a starting point, we apply the decision tree classification algorithm to the entire dataset. To clarify the process further, the classification algorithm is fed with all features as shown in Table 2, and returns the predicted place category. We assume the user's location is the same as each venue's reported coordinates. Therefore, given the user's notification handling behaviour, their location, and the device state, we attempt to predict the type of venue that they are currently at. As seen in Fig. 5, the classification performance is quite good for most categories (F-score macro μ. 82.9%, σ=12.6%). During analysis, we noted that there is some discrepancy in the confidence reported for the most likely current user place, across the place categories (Fig. 6). For this reason, we decided to repeat the analysis in multiple steps, each time limiting the dataset to contain only notifications reported where the most likely current user place was reported above a certain confidence threshold $T \in [0, 0.1, ..0.9]$. The results are shown in Fig. 7. We note that the average F-score is not majorly affected by the reduction of the dataset, however the best nominal performance is achieved when considering venues reported with a confidence threshold $T \geq 0.7$ (μ=84.6%, σ=13.51%, dataset size = 13,558 entries).

4.3 Modelling with inaccurate user coordinates

In the preceding analysis, we assumed that a user's current coordinates are the same as those corresponding to places reported by Google's API. Of course, it would be rare that the user's actual coordinates would be precisely the same as those that match a specific venue, especially for venues that cover a large area (e.g. outdoor parks). To overcome this limitation, we proceeded to modify the user's coordinates by adding random noise to the known place coordinates (latitude and longitude). This noise was applied to each coordinate component individually, following a Gaussian distribution with a standard deviation set by us. The noise standard deviation was calculated using the formula $n \times 10^{-x}$ and was applied to each coordinate component (latitude and longitude), therefore the resulting random coordinates would fall within a certain circular range of a specific venue. An example of how this process generates the random user coordinates within a gaussian distance distribution of a specific venue is shown in Table 4. Distance is calculated using the Haversine formula.

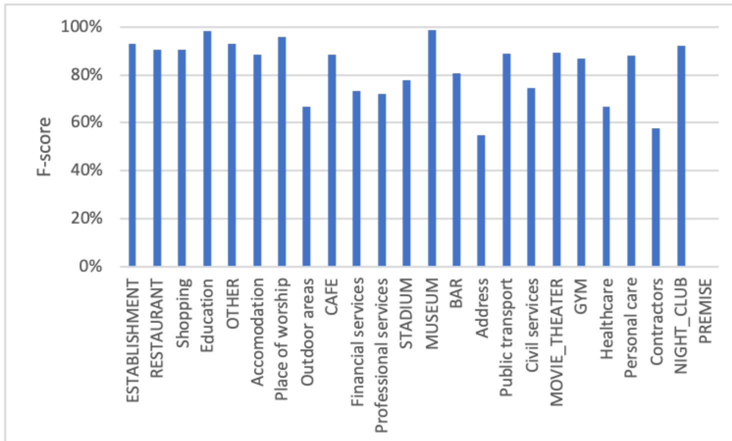


Fig. 5: Average F-score using decision trees, all notifications

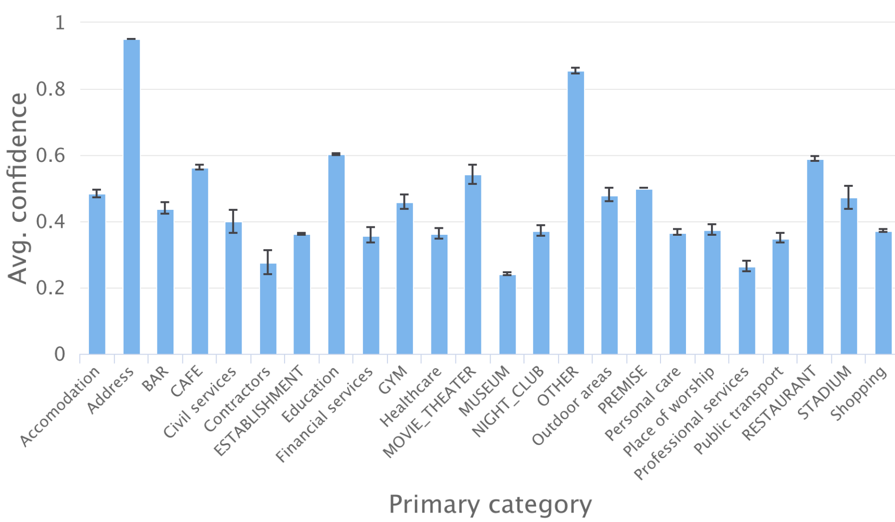


Fig. 6: Average confidence of most likely user place, all notifications, error bars at 95%c.i.

Table 4: Sample random coordinate range generation

Noise σ	**Lat**	**Lng**	**Dist. at 1σ (m)**
0 (Place coords.)	38.2836678	21.7889705	0
1.0×10^{-6}	38.28370608	21.78899229	4.7
2.0×10^{-6}	38.28374437	21.78901408	9.3
3.0×10^{-6}	38.28378265	21.78903587	14.0
4.0×10^{-6}	38.28382093	21.78905766	18.6
5.0×10^{-6}	38.28385922	21.78907944	23.3

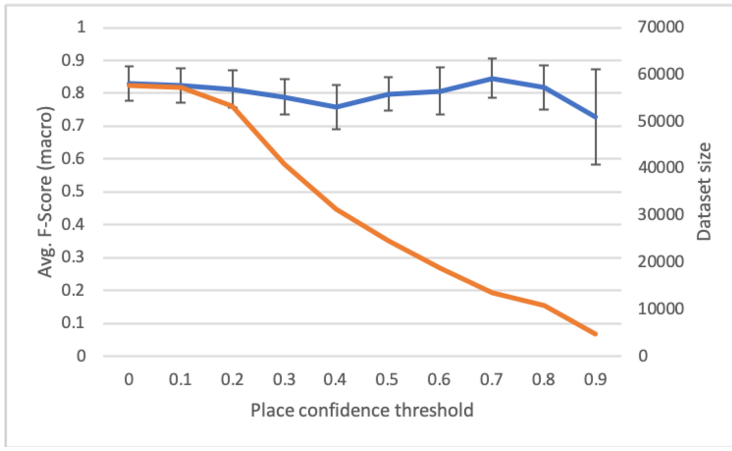
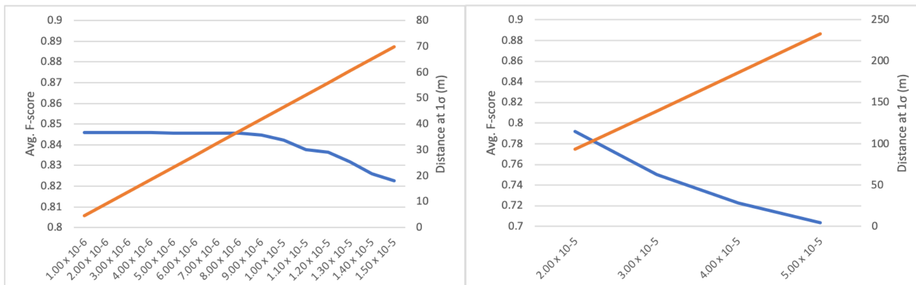


Fig. 7: Average F-score using decision trees (error bars at 95%c.i.)

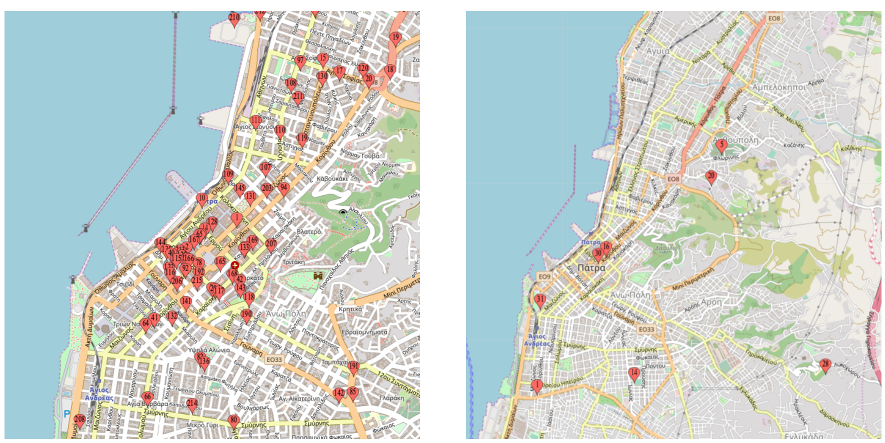
To assess the effect of imprecise user coordinates, we repeated the analysis for each value of $n \in [1, 2, .., 9]$, limiting the dataset to locations with a confidence threshold $T \geq 0.7$, since this achieved the best nominal performance in the preceding analysis. As can be seen in Fig. 8a, the algorithm remains quite robust when adding noise to the decimal coordinates with a $\sigma \leq 9 \times 10^{-6}$ (≈42m), after which, performance begins to deteriorate significantly. We performed also the process for a few larger distances (Fig. 8b). As expected, the performance degradation continues.

At this point, it becomes interesting to observe which categories suffer the heaviest penalty then the user coordinates are further away from the actual place coordinates. Taking the largest noise σ distance (233.1m), we note that the categories Place of worship, Outdoor areas, Professional services, Stadium and Civil services take the worst hit between -35.46% and -55.45% reduction of their F-score, compared to the smallest σ (4.7m). On the other hand, some categories like Shopping and Cafe only take a small penalty (-7.56% and -7.20%) respectively. The explanation for this is possibly rests in the spatial clustering of these venue types (e.g. see Fig. 9). In Fig. 9, we see that cafes are mostly clustered together, hence we may not be able to accurately guess *exactly which* cafe a user is at, but we can be quite certain that they might be at *some* cafe, as long as their location and notification response behaviour is proximal to that captured at a nearby cafes. Although this might suggest that spatial distribution may have a significant effect on the accuracy of the classifier, it must be borne in mind that this is a very extreme scenario. Most users' location data is obtained via A-GPS, which, in an urban environment, has been shown to have an accuracy of about 9m [14].



(a) $4.7m \leq \sigma \leq 69.9m$ (b) $93.2m \leq \sigma \leq 233.1m$

Fig. 8: Average F-score using decision trees, under random coordinate input noise



(a) Cafe category (b) Places of worship category

Fig. 9: Spatial distribution of places in our dataset

4.4 Effect of user location coordinates

In the preceding analysis, one of the input features is the user's location. This feature is certainly obtainable from the user, but its availability depends on whether a user has enabled positioning on their device, their surroundings (indoors or outdoors) and connectivity (wi-fi, 4G, off). So far we have demonstrated that guessing the user's current location type is possible based on their notification behaviour, device state and geographic location, even if the latter is not precisely correspondent to a known place. For the next step, we wanted to experiment without taking user position coordinates into account. The same process as in the previous analysis was repeated, limiting the dataset iteratively to contain notifications at locations above a confidence threshold T. As shown in Fig. 10 the results are much worse than in our previous analysis, showing that the prediction model depends heavily on the knowledge of the user's coordinates, even though these do not necessarily need to correspond with great precision to the true location's coordinates.

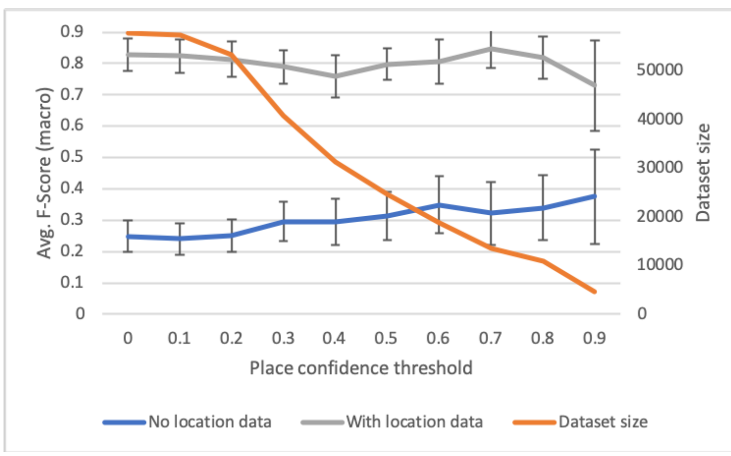


Fig. 10: Average F-score using decision trees (error bars at 95%c.i.)

4.5 Discussion

In this paper, we examined the use of notification handling behaviour as a cue for semantically labelling the user's current location. We found that, when paired with location coordinates, the resulting models can yield useful results with high classification accuracy. Such models can be pre-trained on the cloud and then stored and ran locally on the user's device, as part of an application or service framework, without the need for an internet connection. Further, we demonstrated that such models are robust to small deviations of user coordinates from the actual place coordinates, thereby allowing for positioning errors, or even, the obfuscation of precise user coordinates, in order to maintain privacy.

An underlying assumption in our analysis is that the user is currently positioned and has a certain non-trivial stay time at the location where the notification was received. This is likely true for most cases - users spend more time stationary at various places, than being mobile. However, further work here could include filtering of notification events during transit times, which in our case could not be done (since we did not keep GPS logs for privacy).

For all preceding analyses, there is another underlying assumption, which is that Google's labelling of the place categories has been used as the ground truth. However, this is not necessarily true. Since many of these places are added by users (and ostensibly curated by a few moderators), the assignment of place categories is not necessarily precise. While for some venue categories this issue might be less pronounced, categories such as "Establishment" are quite vague and therefore likely to contain many inaccuracies. As an example, our dataset contains 513 distinct places of type "Establishment". A manual search of this list reveals that 33 of these venues would be better classified under the type "Education" (e.g. "Department of Civil Engineering"). Therefore, it must be noted that as with other studies that leverage social network data (e.g. [5]) the algorithms are tuned to predict the ground truth as reported by the location identification services, therefore introducing an inherent element of inaccuracy. In future work, it would be interesting to examine the failings of these classification algorithms, although this would require a significant investment of time and effort to obtain a reliable ground truth. However, where such algorithms fail, there might be an opportunity to exploit these failures in order to flag venues with incorrect labelling, helping thus to better curate such location datasets.

Finally, as we note different behaviours across venue categories, it would be of value to learn the reasons leading to these variances in user behaviour. However, this would be the subject of a further qualitative study. The generalisability of the findings presented here is limited to the body of the participants (students), hence the varying distribution of sample across categories. The models can be improved by mining information from other populations, to build up the number of samples across as many categories as possible. Personalised models depending on user type can then also be applied to better improve classification performance.

References

1. Akosa, J.S.: Predictive Accuracy : A Misleading Performance Measure for Highly Imbalanced Data
2. Anderson, C., Hübener, I., Seipp, A.K., Ohly, S., David, K., Pejovic, V.: A Survey of Attention Management Systems in Ubiquitous Computing Environments **2**(2), 58:1–58:27. https://doi.org/10.1145/3214261
3. Auda, J., Weber, D., Voit, A., Schneegass, S.: Understanding User Preferences Towards Rule-based Notification Deferral. In: Extended Abstracts of the 2018 CHI

Conference on Human Factors in Computing Systems. pp. LBW584:1–LBW584:6. CHI EA '18, ACM. https://doi.org/10.1145/3170427.3188688

4. Celik, S.C., Incel, O.D.: Semantic place prediction from crowd-sensed mobile phone data **9**(6), 2109–2124. https://doi.org/10.1007/s12652-017-0549-6
5. Falcone, D., Mascolo, C., Comito, C., Talia, D., Crowcroft, J.: What is this place? Inferring place categories through user patterns identification in geo-tagged tweets. In: 6th International Conference on Mobile Computing, Applications and Services. pp. 10–19. https://doi.org/10.4108/icst.mobicase.2014.257683
6. Gu, Y., Yao, Y., Liu, W., Song, J.: We Know Where You Are: Home Location Identification in Location-Based Social Networks. In: 2016 25th International Conference on Computer Communication and Networks (ICCCN). pp. 1–9. https://doi.org/10.1109/ICCCN.2016.7568598
7. He, T., Yin, H., Chen, Z., Zhou, X., Sadiq, S., Luo, B.: A Spatial-Temporal Topic Model for the Semantic Annotation of POIs in LBSNs **8**(1), 12:1–12:24. https://doi.org/10.1145/2905373
8. Komninos, A., Frengkou, E., Garofalakis, J.: Predicting User Responsiveness to Smartphone Notifications for Edge Computing. In: Kameas, A., Stathis, K. (eds.) Ambient Intelligence. pp. 3–19. Lecture Notes in Computer Science, Springer International Publishing
9. Krumm, J., Rouhana, D.: Placer: Semantic Place Labels from Diary Data. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 163–172. UbiComp '13, ACM. https://doi.org/10.1145/2493432.2493504
10. Leppäkoski, H., Rivero-Rodriguez, A., Rautalin, S., Muñoz Martínez, D., Käppi, J., Ali-Löytty, S., Piché, R.: Semantic Labeling of User Location Context Based on Phone Usage Features. https://doi.org/10.1155/2017/3876906, https://www.hindawi.com/journals/misy/2017/3876906/
11. Saikia, P., She, J.: Effective Mobile Notification Recommendation Using Social Nature of Locations. In: 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech). pp. 1265–1270. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.203
12. Visuri, A., van Berkel, N., Okoshi, T., Goncalves, J., Kostakos, V.: Understanding smartphone notifications' user interactions and content importance **128**, 72–85. https://doi.org/10.1016/j.ijhcs.2019.03.001
13. Wu, X., Chen, L., Lv, M., Han, M., Chen, G.: Cost-Sensitive Semi-Supervised Personalized Semantic Place Label Recognition Using Multi-Context Data **1**(3), 116:1–116:14. https://doi.org/10.1145/3131903
14. Zandbergen, P.A.: Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFi and Cellular Positioning **13**(s1), 5–25. https://doi.org/10.1111/j.1467-9671.2009.01152.x
15. Zhu, Y., Zhong, E., Lu, Z., Yang, Q.: Feature engineering for semantic place prediction **9**(6), 772–783. https://doi.org/10.1016/j.pmcj.2013.07.004