



Discovering Daily POI Exploitation Using LTE Cell Tower Access Traces in Urban Environment

Sumin Han¹, Kinam Park¹, and Dongman Lee¹✉

School of Computing, KAIST, Daejeon, South Korea
{suminhan, parkinam321, dlee}@kaist.ac.kr

Abstract. Point of interest (POI) in an urban space represents the perception of city dwellers and visitors of a certain place. LTE cell tower access trace data is one of the promising data sources which has the potential to show real-time POI exploitation analysis. However, there is not much discussion on how it is correlated to diachronic POIs and their exploitation pattern. In this paper, we first show that the access trace pattern from the LTE cell tower can be used to discover which types of POIs exist in a certain area. Then, we propose a daily POI exploitation discovery scheme which can extract patterns of how POIs are daily used. Our analysis can provide a good insight into future urban space-based services such as urban planning and tourism.

Keywords: LTE cell tower access trace · Place of interest · POI exploitation · Classification · Regression

1 Introduction

Understanding how urban spaces are exploited by people can change the way we perceive and conceptualize the city. Service providers can use it to provide citizens with modern urban services such as tourism or personalized advertisement. We can discover such place exploitation based not only on the physical form of a place but also on the projection of human perception. The former is composed of a space and its buildings, while the latter is built from the social and physical perceptions of dwellers accumulated over a long period [1]. Among various forms of data sources to understand the place exploitation [2–6], LTE cell tower access trace data has recently attracted the attention of urban researchers as a promising data source. It can capture the mobility patterns of urban users, which may represent fundamental aspects of social and economic interactions. This advantage with the widespread adoption of mobile devices enables urban researchers to conduct various spatio-temporal analysis without extensive fieldwork. Discovering the correlation between the LTE cell tower access traces and the urban circumstances can broaden the understanding of the placeness of an area, and have the potential to make the real-time place analysis possible as they can provide more instant observation about an urban area.

Liu et al. [6] cluster the moving patterns of citizens read from the mobile traffic data to discover different types of urban functional areas. Zinman et al. [8] identify the land use of urban areas using a classification method on cellular communication usage patterns. Xu et al. [10] cluster the mobile traffic patterns from LTE cell towers to identify each cluster’s urban function recognized by the distribution of POIs. Cici et al. [5] use a clustering approach on mobile traffic data to discover the geographical relationship between the clusters and the POIs in the urban area. Although these clustering and classification methods are informative to a macroscopic urban design, they are limited to discover microscopic exploitation of places and their daily changes, which can provide more temporal information for urban services such as urban tourism and popular place discovery.

In this paper, we propose a POI exploitation discovery scheme that infers which POIs diachronically exist as well as how they are daily exploited in a given area, using LTE cell tower access traces. To analyze spatio-temporal characteristics of urban places, we apply time decomposition to LTE access trace collected from 4,096 LTE cell towers and extract 7-day (168-h) seasonal components. We leverage several classification models that can identify urban space POIs, and regression models for discovering daily POI exploitation, respectively, on the data. For ground truth for POI identification, we annotate each LTE cell using the places and their types from the Google Place API. We also annotate the daily exploitation patterns of each POI in an LTE cell using Google Place popular-times data. Based on these two datasets, our scheme can discover the types of diachronic POIs and the daily exploitation patterns of POIs of an LTE cell area. Evaluation results show that the proposed scheme performs well on the POI identification and can discover explainable changes in POI exploitation.

2 Related Works

Liu et al. [6] use CDR (Call Daily Records) data to discover the fabric of urban city. They extract snapshot map to represent each region and leverage embedded features from the convolutional auto encoder trained by their dataset. Although their evaluation was limited due to the lack of data size, each cluster was shown to include same type of urban functional areas.

Zinman et al. [8] use CDR data to identify the land use of an area which also includes road information such as highway and street. They extract several different features and measure the variable importance (VI) of each feature to make better input feature for the model. They apply random forest classification and their result shows that their analysis of the feature improved accuracy.

Xu et al. [10] extract and model the traffic patterns of large scale cell towers placed in an urban area. They apply hierarchical clustering to seasonal components of mobile LTE traffic data and find a correlation between the cluster and urban ecology. Besides, through traffic spectrum analysis in the frequency domain using DFT, they find a linear combination that could express human activity behavior and indirectly capture the movement of people over time.

Zhang et al. [9] aim to understand the patterns of mobile traffic data and to find correlations between mobile traffic patterns and human activities in urban environments. They separate traffic data into three components: seasonal, trend, and residual. The seasonal component can identify the relationship with the POI of the place through clustering, and the trend component can identify some POI through the pattern difference between weekday and weekend. Residual components show to enable the capture of unexpected behaviors.

Cici et al. [5] aim to infer the characteristics of urban ecology (social and economic activities, and social interaction) from the patterns of cell phone data. They divide the cell phone traffic data in Milan into seasonal and residual patterns using Fourier time decomposition. They find the seasonal data characterize the patterns related to socio-economic activity in the region using agglomerative hierarchical clustering, while residual data make it possible to analyze how irregularity from the new events affect in other regions.

Shafiq et al. [7] aim to identify application usage patterns in each region so that network operators can easily optimize the distribution of cellular data. They collect traffic volume in terms of byte, packet, flow count, and unique user count, which can be extracted from 3G cellular data, and clustered for each term to specify the characteristics of each region (Downtown, University, Suburb, etc). The network operators can optimize the distribution of cellular data that fits clusters.

3 Dataset

In this section, we describe how we construct our dataset in three steps. We first explain how we preprocess LTE signal patterns. Then we illustrate how we annotate each LTE cell with POI type label, and collect POI exploitation patterns for the experiment.

3.1 Step1: Extracting a 7-Day Seasonal Component from an LTE Cell Tower Access Trace

Our dataset consists of anonymized LTE cell tower access traces on 4,096 LTE cell towers, collected by an internet service provider (ISP), Korea Telecom (KT) from March 1st, 2018 to February 28th, 2019. These LTE cell towers are located across three most popular commercial areas in Seoul as shown in Fig. 1. Each cell tower covers a $50\text{ m} \times 50\text{ m}$ area and access traces represent the number of people connected to each cell per hour. To analyze embedded patterns from each LTE cell tower access trace, we leverage the STL time decomposition method [22] to extract a 7-day (168-h) seasonal components. Then we conduct min-max scaling to normalize each pattern.

3.2 Step 2: Associating Diachronic POI Type Labels with Each Cell

To understand which types of diachronic POIs exist in each LTE cell, we establish a ground-truth of (LTE access pattern, POI labels) pairs. First of all, we crawl

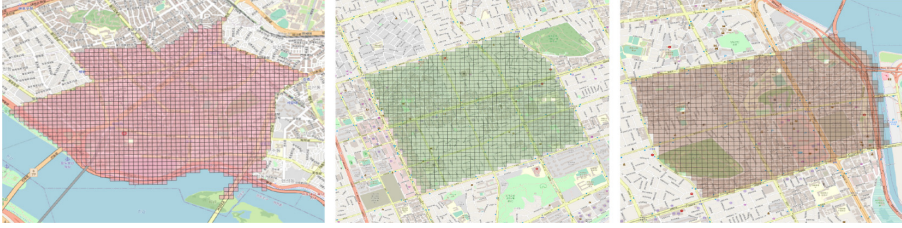


Fig. 1. We use LTE cell tower access traces from $50\text{ m} \times 50\text{ m}$ LTE cells in the described area. Hongik-University (Area 1, left-most: $\text{N}37.551^\circ$, $\text{E}126.925^\circ$), Yeoksam-station (Area 2, center: $\text{N}37.501^\circ$, $\text{E}127.036^\circ$), and Samseong-station (Area 3, right-most: $\text{N}37.514^\circ$, $\text{E}127.052^\circ$).

a list of places in all the LTE cells using Google Place API¹ by querying any places within a radius of 36 m from each centroid of the LTE cells to be able to cover each cell area. Then we use the place type tags from Google query result to annotate LTE cell whether it contains certain types of POI. To prevent the associations of the POI labels with the LTE cell tower access patterns from being diversified too much, we choose 8 most frequent place types (out of 100) found by the Google Place API in our three target areas: *restaurant*, *store*, *cafe*, *health*², *finance*, *bar*, *clothing_store*, and *lodging*. Table 1 shows the number of LTE cells which contains places of corresponding POI types. Figure 2 shows the median of the extracted 7-day seasonal component for each type. We can see a clear difference between weekdays and weekends. Weekdays have regular patterns, while weekends have irregular patterns.

Table 1. The number of LTE cells containing each POI type.

	restaurant	store	cafe	health	finance	bar	clothing_store	lodging
Area 1	1,953	1,503	913	378	126	449	346	337
Area 2	1,317	1,009	417	651	418	218	85	172
Area 3	623	683	181	277	173	102	154	53
Total	3,893	3,195	1,511	1,306	717	769	585	562

3.3 Step 3: Annotating Daily POI Exploitation Using Google Popular-Times

To understand how the exploitation patterns can be inferred from LTE access pattern, we establish a ground-truth of (LTE access pattern, a daily POI

¹ <https://developers.google.com/places/web-service/search>.

² This category is mostly comprised of medical clinics.

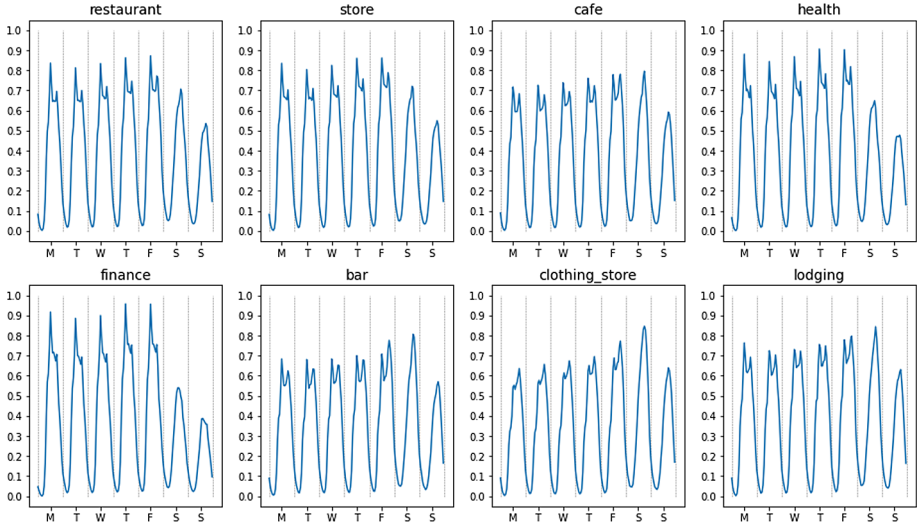


Fig. 2. The median of min-max scaled weekly seasonal patterns from each POI type. Every 24h represents the day of the week starting from Monday to Sunday.

exploitation pattern) pairs. This dataset requires POI exploitation patterns on a daily basis to figure out how POIs in a given cell are exploited during a week. To construct such a dataset, we use Google popular-times³ for each POI. Since official Google Place API does not provide popular-times, we leverage the third-party crawler⁴ that provides a 7-day exploitation pattern normalized from 0 to 100 for a given place. We simply scale the pattern by 1/100 to have it represented by values between 0 and 1. Table 2 shows the number of collected POI exploitation data in each area.

Table 2. The number of POIs contain popular-times information.

	restaurant	store	cafe	health	finance	bar	clothing_store	lodging
Area 1	541	193	260	2	19	108	19	0
Area 2	323	105	65	6	26	38	13	3
Area 3	133	71	29	1	14	16	13	2
Total	997	369	354	9	59	162	45	5

³ <https://support.google.com/business/answer/6263531?hl=en>.

⁴ <https://github.com/m-wrzt/populartimes>.

4 Methods

4.1 POI Discovery

To discover POIs in each LTE cell from the access trace pattern, we leverage logistic classifier (LC), support vector machine (SVM), random forest classifier (RFC), and deep neural network (DNN) based model. For each POI type, each model listed above is trained to predict whether the LTE cell area may contain target POI type or not, which can be considered as a binary classification problem. The DNN based model consists of 3 linear layers with ReLU activation function on two foremost layers, and a final sigmoid function for binary classification. This model produces the prediction score from 0 to 1 where a higher value implies that a target POI type is more likely to exist in a given LTE cell. The detailed architecture is illustrated in Fig. 3.

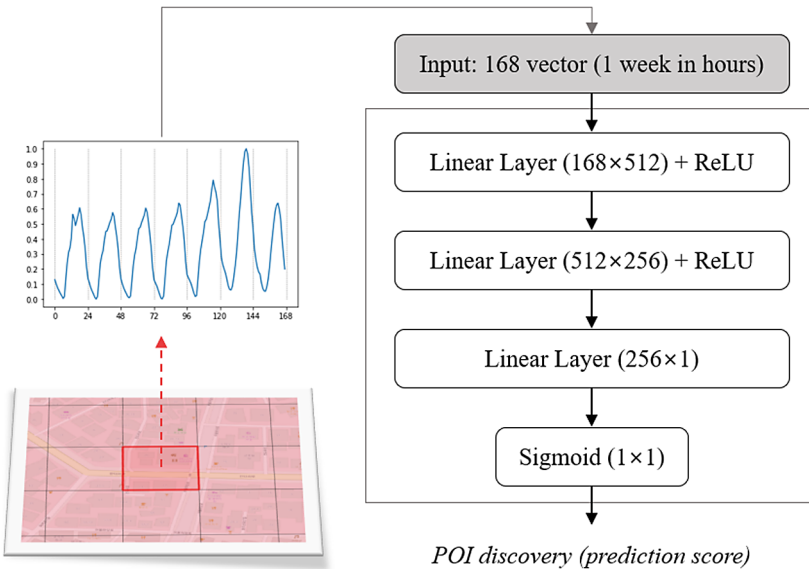


Fig. 3. The DNN based POI discovery model architecture.

4.2 Daily POI Exploitation Pattern Discovery

To discover the daily POI exploitation pattern from an LTE access pattern, we leverage a random forest regression (RFR) model which can be used for time-series pattern prediction [20, 21]. Figure 4 illustrates our RFR model and its training process. First of all, find POIs with exploitation patterns for each LTE cell. Next, pair each POI exploitation pattern with corresponding LTE

access pattern. During this process, pairs of an LTE access pattern and a POI exploitation pattern will be separated into different datasets each belongs to the same POI type. Then, we train a RFR model to predict the exploitation pattern for each POI type. Finally, this model can be used to predict unknown POI exploitation pattern using corresponding LTE access pattern in a given area.

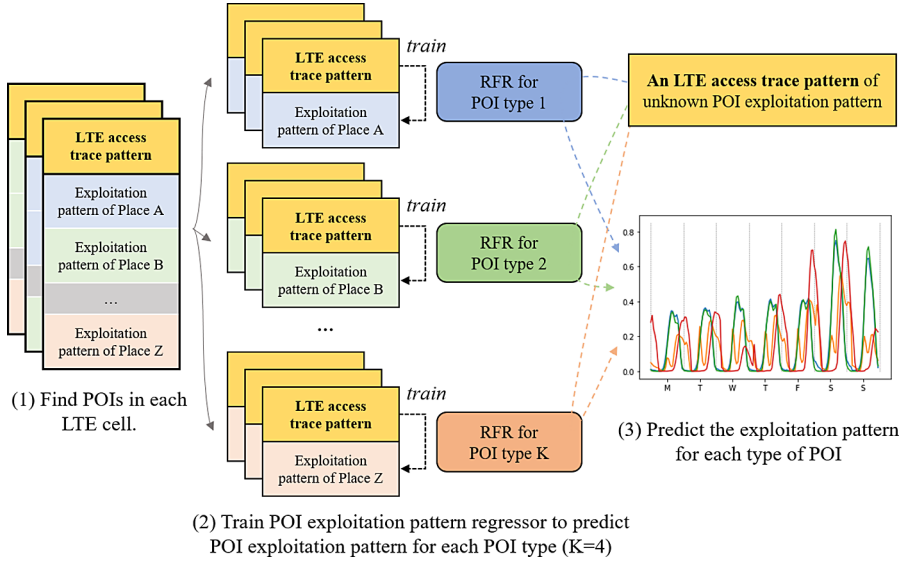


Fig. 4. Daily POI exploitation discovery scheme on an LTE access trace pattern.

5 Evaluation

5.1 Evaluation Setup

POI Discovery. We train binary classification models (LC, SVM, RFC, DNN model) for each POI type to predict whether the cell contains such POI or not. When we train the models, they may suffer from the unbalanced amount of positive and negative labels of the dataset. To prevent this, we use a random under-sampling technique that removes either positive or negative samples randomly to have an equal amount to each other. We conduct 10-fold cross-validation using the dataset described in Sect. 3.2. To analyze the performance of the model, we use five standard metrics: (1) *Accuracy*, (2) *Precision*, (3) *Recall*, (4) *F1-macro*, (5) *ROC-AUC*. When we evaluate the first four metrics above, we use the threshold value as 0.5 to make a positive prediction. For each metric in this experiment, a higher value represents that a model shows better performance. The parameters used in each model are as follows:

- LC: L2 norm loss, tolerance = 0.1, LBFGS solver
- SVM: radial basis function (RBF) kernel, regularization parameter = 1
- RFC: the maximum depth of 5, the number of estimators = 100

Table 3 shows the mean value for each metric after cross validation with respect to each POI type.

Daily POI Exploitation Discovery. We train random forest (RFR) regression models with maximum-depth of 5. We compare the performance of RFR models to the baseline which uses the mean and median of the training data to predict the test data. To evaluate our model, we conduct 10-fold cross-validation using the dataset described in Sect. 3.3. During the experiment, we find that many places have one or two days of holiday every week. However, this can cause a lot of error on evaluation if the model predicts any activated pattern on the holiday of a POI. We omit the activated signals on the holiday of the POI from the model prediction by zeroing them. We simply define a holiday of a POI as a day whose corresponding 24-h pattern has the values equal to 0. We use three evaluation metrics, which are (1) *mean squared error (MSE)*, (2) *root mean squared error (RMSE)*, and (3) *Pearson correlation*. For MSE and RMSE, a smaller value represents that the model produces a less amount of errors. For the Pearson correlation, a higher value represents that the predicted pattern is more similar to the corresponding real POI exploitation pattern. We also evaluate Pearson correlation by mean, median, standard deviation (SD), and coefficient of variation (CV) between the test results. We choose a maximum depth of 5 for an RFR model.

5.2 Validation of POI Discovery

Table 3 shows the performance results of each model for each POI type. When we compare the overall performance across the categories, the models trained for bar category show the highest accuracy while that for the health category show the lowest. Although the accuracy of a model for bar might be seemed marginally better than others, the precision is higher than any other category. This is because the signal patterns of LTE cells encompassing bars are quite distinguishable from others as bars mainly open during the night and their existence are easy to be discovered from the patterns. On the other hand, POIs in the health category are difficult to be predicted as its corresponding places do have a small number of visitors, compared with other categories. This makes it difficult to find a common pattern from the LTE access pattern belonging to this category. The table shows that RFC and DNN model can be considered as a model which produces a fine prediction in terms of accuracy, as they show good performance on all the categories except for clothing.store and lodging. LC and SVM also performs pretty well and worth to be applied when the researcher needs an prototype result. However, evaluating the models only on accuracy can be dangerous as we conduct our experiment on a small amount of dataset after

Table 3. The evaluated performance measures for diachronic POI discovery.

POI	Model	Accuracy	Precision	Recall	F1-macro	AUC-ROC
restaurant	LC	0.6881	0.6789	0.7107	0.6940	0.7567
	SVM	0.7072	0.6969	0.7301	0.7127	0.7675
	RFC	0.6858	0.6704	0.7331	0.6990	0.7643
	DNN	0.7093	0.7076	0.7442	0.7159	0.7828
store	LC	0.6566	0.6447	0.6992	0.6697	0.7316
	SVM	0.6771	0.6643	0.7204	0.6896	0.7447
	RFC	0.6987	0.6917	0.7150	0.7025	0.7730
	DNN	0.6911	0.6972	0.7565	0.7062	0.7614
cafe	LC	0.6902	0.6958	0.6774	0.6849	0.6903
	SVM	0.6952	0.7053	0.6724	0.6867	0.6946
	RFC	0.6957	0.6925	0.6996	0.6948	0.7682
	DNN	0.6979	0.6938	0.7243	0.6994	0.7648
health	LC	0.6181	0.5998	0.7159	0.6507	0.6457
	SVM	0.6160	0.5946	0.7244	0.6520	0.6587
	RFC	0.6368	0.6237	0.6910	0.6548	0.6844
	DNN	0.6271	0.6185	0.7121	0.6555	0.6685
finance	LC	0.6617	0.6463	0.7186	0.6787	0.6603
	SVM	0.6878	0.6684	0.7360	0.6992	0.6852
	RFC	0.6992	0.6960	0.7044	0.6990	0.7459
	DNN	0.7013	0.7027	0.7420	0.7077	0.7477
bar	LC	0.7204	0.7264	0.7052	0.7138	0.7219
	SVM	0.7320	0.7594	0.6728	0.7119	0.7335
	RFC	0.7278	0.7271	0.7185	0.7212	0.7998
	DNN	0.7384	0.7404	0.7417	0.7304	0.8068
clothing_store	LC	0.6677	0.6799	0.6330	0.6533	0.6680
	SVM	0.6895	0.7272	0.6127	0.6612	0.6896
	RFC	0.6664	0.6855	0.6249	0.6502	0.7272
	DNN	0.6581	0.6705	0.6607	0.6582	0.7396
lodging	LC	0.6830	0.6764	0.7009	0.6857	0.6822
	SVM	0.6925	0.7020	0.6656	0.6815	0.6922
	RFC	0.6845	0.6818	0.6909	0.6849	0.7246
	DNN	0.6736	0.6662	0.7290	0.6870	0.7314

the under-sampling technique. So, we also evaluate the models based on AUC-ROC scores, which can take into account of sensitivity and specificity. In terms of AUC-ROC scores, DNN based model shows better performance overall. We

assume that this is because the DNN architecture captures a feature that most commonly appears in each POI type, resulting in a more consistent model.

Table 4. The performance measures of POI exploitation pattern prediction.

POI	Model	MSE	RMSE	Pearson correlation			
				Mean	Median	SD	CV
restaurant	Mean	0.0313	0.1769	0.6979	0.7264	0.1636	0.2346
	Median	0.0336	0.1832	0.6832	0.7156	0.1793	0.2627
	RFR	0.0274	0.1655	0.7459	0.7827	0.1702	0.2284
store	Mean	0.0336	0.1831	0.7486	0.8137	0.2092	0.2809
	Median	0.0353	0.1877	0.7414	0.8207	0.2286	0.3101
	RFR	0.0310	0.1761	0.7685	0.8303	0.2000	0.2615
cafe	Mean	0.0286	0.1690	0.8000	0.8450	0.1482	0.1857
	Median	0.0303	0.1739	0.7922	0.8437	0.1591	0.2013
	RFR	0.0252	0.1585	0.8264	0.8699	0.1412	0.1713
bar	Mean	0.0250	0.1580	0.7569	0.8353	0.2229	0.2967
	Median	0.0267	0.1632	0.7496	0.8336	0.2334	0.3137
	RFR	0.0231	0.1519	0.7731	0.8579	0.2222	0.2892

5.3 Validation of Daily POI Exploitation Discovery

We conduct our experiment on *restaurant*, *store*, *cafe*, and *bar* since these categories have enough data to evaluate our experiment while others suffer from data sparsity. Table 4 shows the evaluated performance measures of daily POI exploitation pattern prediction. The results show that RFR-based models produce best prediction patterns than baseline mean and median models as they show less MSE and RMSE. The models for bar category show lowest MSE and RMSE as the exploitation patterns of this category are more distinguishable than others. The higher mean value of Pearson correlations means that RFR models predict rather similar exploitation patterns to the real value than the baseline models. Plus, a small coefficient of variance means that the level of dispersion of the Pearson Correlation is smaller around the mean. Therefore, we can conclude that RFR models produce a similar exploitation pattern to the corresponding real value more frequently for all types of POI. In addition, a gap between the mean of the Pearson correlation from the best RFR model and that from the baseline mean model is biggest in the restaurant category and smallest in the bar category as shown in Table 4. Considering that the amount of training data is biggest at the restaurant category and smallest at the bar category, this observation implies that the current RFR model can be reinforced as more training data is provided.



Fig. 5. Different daily changes of POI exploitation of 4 places with the same POI set derived from the LTE access pattern in Area 1. The colored circle on the map represents each area: *restaurant* (blue), *store* (orange), *cafe* (green), and *bar* (red). We emphasize the dominated POI pattern for each POI exploitation graph by thickening the line width. (Color figure online)

5.4 Finding POI Exploitation Variances Across Places with the Same POI Types

By conducting a cross-analysis of the daily POI exploitation changes, it can be found that urban places with the same POI types can be exploited differently by dwellers and tourists from the ne-grained perspective. Figure 5 shows the daily changes of POI exploitation of four different places identified by the same diachronic POIs (i.e., having similar urban compositions), while their dominant POI types are different. By using the Google street view, we find that exploitation patterns are different, depending on the urban setting of each place. A restaurant dominant place (blue) is located near a main street where people gather and socialize while having lunch and dinner. Its exploitation pattern shows that this area is highly visited at every lunch and dinner periods, and even higher at the dinner period on Friday and Saturday. A store dominant place (orange) is a popular shopping street mainly for buying clothes and accessories. The

exploitation pattern of this place shows that people visit during all the day, unlike the restaurant dominated place, which shows clear peaks in lunch and dinner periods. A cafe dominant place (*green*) includes big franchise cafes (e.g., Starbucks) and many unique theme cafes (pet, flower, movie character, etc.). Its exploitation pattern shows that people visit the place during daytime, and more on a weekend as there are many special type of cafes which are attractive to visitors. A bar dominant place (*red*) is well-known for bars, night clubs, concerts, and night food. Its exploitation pattern shows that people visit the place over midnight, and more on Friday and Saturday. It is interesting to note that the exploitation pattern of restaurants in this place is similar to that of bars, that is, many restaurants in this area are also not closed even after midnight. The findings above show that our proposed scheme can provide urban planners who seek for static configuration and dynamic exploitation of urban places [16, 17], and urban tourists who want to visit places that match with their daily interest with helpful insights [18].

5.5 Discussion

The results above show that we can discover POIs and their daily exploitation leveraging several classification and regression models using a normalized 7-day seasonal pattern from an LTE access traces. Nevertheless, there still exist a few points to investigate further. First, we use a normalized 7-day seasonal pattern from each LTE access pattern as our input for both experiments. However, during the extraction of seasonal patterns, many information such as trend or residual patterns are missed. A trend pattern can reflect one-year seasonal change such as the weather or public traveling behavior, and it can provide more insight on the long-term exploitation pattern change. A residual pattern can reflect unexpected signal away from seasonal and trend, so it can show the frequency of unexpected event in a given area. So, the combined analysis of trend and residual as well as seasonal patterns would provide another dimension to analyze the LTE access pattern. Another way to read social information from LTE access patterns is to analyze in terms of age and gender. This will enable us to understand the demographics of the area and help to figure out what age and gender groups facilitate the POI (or a combination of POIs) of a given place - that is, understand what placeness is leveraged not only when but also by whom.

6 Conclusion

In this paper, we leverage LTE cell tower access traces and develop a decent scheme to identify what kinds of points of interest (POIs) exhibit in a given space and how these POIs are exploited in a daily manner. For the former, we leverage several binary classification models such as LC, SVM, RFC, and DNN. For the latter, we leverage the random forest regression model for discovering daily changes in POI exploitation. Evaluation results show that our approach performs well on each task, and can discover the poi exploitation variance of the

different urban environment. Our analysis based on LTE cell tower access data will provide a good insight into future urban space-based services such as travel and tourism.

Acknowledgment. This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2019-0-01126, Self-learning based Autonomic IoT Edge Computing). We appreciate Korea Telecom(KT) for providing valuable LTE access data.

References

1. Loukaitou-Sideris, A., Banerjee, T.: *Urban Design Downtown: Poetics and Politics of Form*. University of California Press, Berkeley (1998)
2. Ferrari, L., et al.: Extracting urban patterns from location-based social networks. In: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (2011)
3. Salamon, J., Bello, J.P.: Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Process. Lett.* **24**(3), 279–283 (2017)
4. Ferreira, N., et al.: Visual exploration of big spatio-temporal urban data: a study of New York city taxi trips. *IEEE Trans. Vis. Comput. Graph.* **19**(12), 2149–2158 (2013)
5. Cici, B., et al.: On the decomposition of cell phone activity patterns and their connection with urban ecology. In: *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing* (2015)
6. Liu, C., et al.: Clustering analysis of urban fabric detection based on mobile traffic data. In: *Journal of Physics: Conference Series*, vol. 1453 (2020)
7. Shafiq, M.Z., et al.: Characterizing geospatial dynamics of application usage in a 3G cellular data network. In: *2012 Proceedings IEEE INFOCOM*. IEEE (2012)
8. Zinman, O., Lerner, B.: Utilizing digital traces of mobile phones for understanding social dynamics in urban areas. *Pers. Ubiquit. Comput.* **24**(4), 535–549 (2019). <https://doi.org/10.1007/s00779-019-01318-w>
9. Zhang, M., Xu, F., Li, Y.: Mobile traffic data decomposition for understanding human urban activities. In: *2016 IEEE 13th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE (2016)
10. Xu, F., et al.: Understanding mobile traffic patterns of large scale cellular towers in urban environment. *IEEE/ACM Trans. Netw.* **25**(2), 1147–1161 (2016)
11. Han, D., Kim, J., Kim, J.: Deep pyramidal residual networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017)
12. Shoji, Y., Takahashi, K., Dürst, M.J., Yamamoto, Y., Ohshima, H.: Location2Vec: generating distributed representation of location by using geo-tagged microblog posts. In: Staab, S., Koltsova, O., Ignatov, D.I. (eds.) *SocInfo 2018*. LNCS, vol. 11186, pp. 261–270. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01159-8_25
13. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
14. Mitchell, S.: *Evaluating impacts and defining public perceptions of police body-worn cameras (BWCs)*. Diss. Kent State University (2019)
15. Hair, J.F., Ringle, C.M., Sarstedt, M.: Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. *Long Range Plan.* **46**(1–2), 1–12 (2013)

16. Frias-Martinez, V., Frias-Martinez, E.: Spectral clustering for sensing urban land use using Twitter activity. *Eng. Appl. Artif. Intell.* **35**, 237–245 (2014)
17. Southworth, M., et al.: People in the design of urban places, pp. 461–465 (2012)
18. Karduni, A., et al.: Urban space explorer: a visual analytics system for urban planning. *IEEE Comput. Graph. Appl.* **37**(5), 50–60 (2017)
19. Abbar, S., Zanouada, T., Al-Emadi, N., Zegour, R.: City of the people, for the people: sensing urban dynamics via social media interactions. In: Staab, S., Koltsova, O., Ignatov, D.I. (eds.) *SocInfo 2018*. LNCS, vol. 11186, pp. 3–14. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01159-8_1
20. Wu, H., et al.: Time series analysis of weekly influenza-like illness rate using a one-year period of factors in random forest regression. *Bioscience Trends* (2017)
21. Shirmohammadi-Khorram, N., et al.: A comparison of three data mining time series models in prediction of monthly brucellosis surveillance data. *Zoonoses Public Health* **66**(7), 759–772 (2019)
22. Cleveland, R.B., et al.: STL: a seasonal-trend decomposition. *J. Official Stat.* **6**(1), 3–73 (1990)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

