Title: Extending prediction of parking occupancy to untracked city areas using GIS metadata

Basic description: Several smart cities around the world have begun monitoring parking areas in order to predict free spots and help drivers that are looking for parking. The current results are indeed promising, however this approach is limited by the high cost of sensors that need to be installed throughout the city in order to achieve an accurate prediction rate. This work investigates the extension of forecasted parking information from areas equipped with sensors to areas that are missing them. To this end, similarity values between city neighborhoods will be computed based on background data from geographic information systems. Using the derived similarity values, the adaptation of occupancy rates from monitored- to unmonitored parking areas will be analysed.

Technical details

Parking sensor information: The data from the SFpark project in San Francisco will be used, which offers the richest such datasets to date[1].

Prediction models: A prediction model will first be computed per street block, as this is the unit for which occupancy data is available. Afterwards, a clustering into larger parking areas will be necessary, without compromising accuracy[2], in order to capture a zone with richer GIS metadata. The prediction models will be computed using the state-of-art methods, like neural networks, autoregression models and support vector regression[2][3].

GIS metadata: City entities that provide parking, i.e. schools, kindergartens, hospitals, museums, theaters, shopping centers, restaurants, etc. will be considered, whereas their capacities and/or geographical sizes will be taken into account. Information sources include OpenStreetMap, Google Maps, Bing Maps, etc.

Similarity measures: Upon representing the GIS metadata in form of vectors or sets, functions like Cosine and Jaccard will be applied.

Determining occupancy values for untracked areas: To calculate the occupancy rates for a new parking zone, the prediction models for the top N most similar known areas will be applied under the respective conditions (day of week, time of day, weather, possibly events and traffic). The results will be subsequently multiplied with the corresponding similarity values to arrive at the final scores. Note that since exact numerical computation is necessary, the prediction models will internally work with exact probability rates before eventually arriving at an occupancy class (e.g. low, medium, high).

Evaluation: To test the accuracy of the inferred occupancy values, areas will be selected for which predictions are available directly from installed sensors.

Originality of work: Up to now, the smart parking literature concentrates on capturing data and predicting parking information using static or mobile sensors, by covering the areas for which future rates are calculated[4]. An extrapolation approach as described here, to the best of my knowledge, has not been explored yet.

References

[1] *SFpark project*. `http://sfpark.org`. Accessed: 2017-01-30.

[2] Xiao Chen. *Parking occupancy prediction and pattern analysis*. Tech. rep. Technical report, Stanford University, 2014. Machine Learning Final Projects.

[3] Tooraj Rajabioun and Petros A Ioannou. "On-street and off-street parking availability prediction using multivariate spatiotemporal models". In: *IEEE Transactions on Intelligent Transportation Systems* 16.5 (2015), pp. 2913–2924.

[4] Trista Shuenying Lin. "Smart parking: Network, infrastructure and urban service". PhD thesis. Lyon, INSA, 2015.