

# Crowdsensing Maps of On-Street Parking Spaces

Vladimir Coric  
Department of Computer and  
Information Sciences  
Temple University  
1805 N. Broad St, Philadelphia, PA  
vladimir.coric@temple.edu

Marco Gruteser  
WINLAB  
Rutgers University  
671 Route 1 South  
North Brunswick, NJ  
gruteser@winlab.rutgers.edu

**Abstract**—It has been estimated that traffic congestion costs the world economy hundreds of billions of dollars each year, increases pollution, and has a negative impact on the overall quality of life in metropolitan areas. A significant part of congestion in urban areas is due to vehicles searching for on-street parking. Detailed and accurate on-street parking maps can help drivers easily locate areas with large numbers of legal parking spaces and thus relieve congestion. In this paper, we address the problem of mapping street parking spaces using vehicles’ preinstalled parking sensors. In particular, we focus on identifying legal parking spaces from crowdsourced data, whereas earlier work has largely assumed that such maps of legal spaces are given. We demonstrate that crowdsensing data from vehicle parking sensors can be used to classify on-street areas into legal/illegal parking spaces. Based on more than 2 million data points collected in Highland Park, NJ and downtown Brooklyn, NY areas, we show that on-street parking maps can be estimated with an accuracy of  $\sim 90\%$  using proposed weighted occupancy rate thresholding algorithm.

## I. INTRODUCTION

It has been estimated that traffic congestion costs the world economy hundreds of billions of dollars each year, increases pollution, and has a negative impact on the overall quality of life in metropolitan areas. In order to solve this emerging problem, transportation departments increasingly rely on systems for real-time traffic control and management known as Intelligent Transportation Systems (ITS) [1]. The main goal of ITS systems is to inform travelers about current and future traffic and motivate them to modify travel plans during congested periods and, in doing so, relieve the congestion. However, in heavily populated areas such as downtowns, vehicle routing becomes very challenging. One of the reasons why it is hard to reroute vehicles in downtown areas is due to the fact that a significant part of congestion in these areas is due to parking. Vehicles searching for available on-street parking spaces slow down traffic and in addition pollute the air by emitting large quantities of carbon dioxide. In a study conducted in a central business district in downtown Los Angeles [2], it was shown that vehicles searching for parking in a period of one year created 38 trips around the world, spending 47,000 gallons of gasoline and releasing 730 tons of carbon dioxide.

Maps of garage locations are available to travelers online for all major cities. However, constructing garages is expensive, hence spaces are limited and prices high and walking distances

to the final destination can be long, therefore an important parking option for most travelers is still on-street parking [3]. On the other hand, due to a lack of information about the number of parked cars on the streets and lack of maps of legal/illegal parking spaces, vehicles spend long periods of time searching for empty parking spots. The problem of having an unknown number of cars parked along the streets was addressed by the SF park project [4], in which sensors were buried under the pavement beneath 25% of the slotted parking spaces in the city of San Francisco in order to detect whether the parking spaces were occupied or not. The high price of installing fixed sensors into the pavement motivated the PARKNET project [5] to employ ultrasonic sensors together with GPS units on several vehicles (taxies, police vehicles, etc.) in order to detect the number of cars parked on the streets. However, none of the above projects address the problem of creating accurate and detailed maps of legal parking spaces. Furthermore, the [5] assumes that those maps are available to city authorities to some extent. Manually creating these maps (ex. from Google Maps [6]) can be a very long and tiresome process. Several projects were done on successfully inferring road maps from GPS traces using data mining algorithms [7], [8] but none of them address the question of inferring parking maps. On the other hand, the existence of these maps can be very useful in several applications. For example, travelers will be able to explore if the area where they desire to park possesses a large number of illegal parking areas such as fire hydrants, private garages, or bus zones. If a high demand on-street parking area contains a large number of illegal parking spaces, the probability that vehicles can park in this area is smaller and travelers can change their parking preferences in advance. Furthermore, detailed parking maps can be implemented into GPS devices and help drivers to identify whether the space where they are currently parked is legal or not. For example, if the driver parked a vehicle close to a fire hydrant or a bus zone, the GPS receiver can beep in order to warn the driver that this is not a legal parking space. This way, travelers can avoid getting parking fines or having their vehicle towed because it was irregularly parked. The parking authority can use parking maps to identify areas with a small number of legal parking spaces in order to recommend the building of parking garages in that area. In addition to previous applications, detailed maps of legal/illegal parking

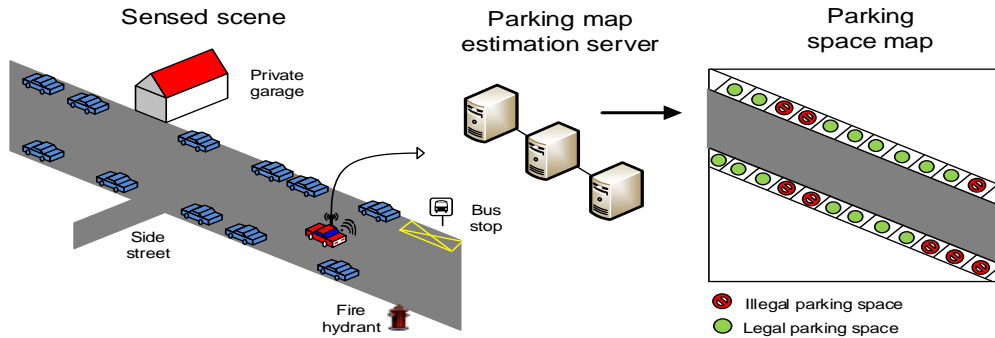


Fig. 1: Overview of the proposed system

spaces can be used in the creation of services similar to [5], where these services will on the one hand report availability of legal parking spots and on the other report to parking authorities if illegal parking spots are occupied.

In this paper, we address the problem of mapping out street parking spaces using car preinstalled parking sensors obtained by crowdsensing. It is important to stress that this project differs from the [5] in that the goal of [5] was to collect space occupancy data as opposed to our project where the main goal is to create legal/illegal parking spot maps, *maps which earlier work assumes already exist and are available*. A large number of new generation vehicles possess range finder parking sensors which help drivers to park their vehicles [9]. While the vehicle is in motion, these sensors can be used to detect the presence or absence of parked vehicles on the street. The sensor measurements can then be reported together with the vehicles GPS coordinates (which can be obtained from the cars preinstalled GPS device) to the centralized server which can estimate if the reported parking spaces are legal or illegal. In order to accurately estimate parking maps, the centralized server would need several measurements of the same location, possibly from several different vehicles. This can be achieved using *crowdsensing*, an approach that collects large amounts of sensing data from crowds. By providing parking sensor measurements, users can help in mapping the streets into legal and illegal parking spaces similarly to other crowdsensing projects where crowdsensing data is used to estimate traffic, monitor pollution levels in a city, estimate bus arrival times and so forth. The diagram of the proposed system is shown in Figure 1.

The rest of the paper is organized as follows. In section 2, we will explain data sets used to demonstrate how parking maps can be created based on ultrasonic parking sensors and give a brief overview of the [5] system which is used to collect parking data. In section 3 we will present algorithms for estimating parking maps from several passes through the same streets and present results of the evaluation of the algorithms in section 4. Finally in section 5 we discuss in more detail some of the issues we encounter followed by related work in section 6 and conclusions in section 7.

## II. ON-STREET PARKING DATA

To demonstrate how on-street parking spaces can be mapped using parking sensors, we used road side parking data obtained from Highland Park, NJ and downtown Brooklyn, NY areas. The data from these two data sets is collected using the [5], a system which collects on-street parking availability information. The system in [5] consists of a low cost ultrasonic sensor (Figure 2a), which measures distance from the car to the nearest obstacle, and the GPS receiver which reports the location of the sensor measurement. The ultrasonic sensor emits ultrasonic waves every 50 ms at the frequency of 43 KHz, providing single range readings from 12 to 255 inches. If an obstacle is detected, the sensor will report the distance from the obstacle to the vehicle and in the case that no obstacle is detected, the sensor will report a distance of 255 inches. The role of the GPS receiver is to provide time stamps and location stamps for each sensor measurement. The collected sensor measurements and corresponding location stamps form time series data which represents passes through the streets. The ultrasonic sensor and GPS receiver were deployed on several vehicles which were cruising around streets and reporting parking measurements to the centralized server where they were preprocessed and used by several algorithms in order to estimate street parking availability. In addition to the ultrasonic sensor and GPS receiver, a small Sony PS3 camera was integrated in order to provide ground truth for parking estimation. The camera was situated just above the sensor and was aligned together with the sensor (Figure 2a). It is important to note that the camera was not part of the [5] system; it was only used for evaluation and data analysis purposes.

### A. Highland Park data set

The first parking data set was collected in 2009 in Highland Park, NJ in three road side parking areas as illustrated in Figure 3a. Three sensing vehicles collected data during a two month period during their daily commutes. They collected more than 500 miles of roadside parking data. The data was collected only from certain streets in Highland Park (Table I) and this was done by using the concept of trip boxes,



(a) Position of ultrasonic sensor and web camera (b) Vehicle parked on the bus stop on Bergen Street (c) Vehicle parked to close to fire hydrant on Schermerhorn Street

Fig. 2: Ultrasonic sensor properties and parking violations

TABLE I: Selected streets for both data sets

Street name	Street length(m)	# of passes	From	To	Bicyc lane
<b>Brooklyn</b>					
<b>Bergen</b>	608	8	Nevins	Smith	no
<b>Smith</b>	311	8	Dean	Scherm	yes
<b>Nevins</b>	382	8	Scherm	Bergen	no
<b>Dean</b>	608	3	Smith	Nevins	no
<b>Scherm.</b>	608	6	Smith	Nevins	yes
<b>Wychoff</b>	608	3	Smith	Nevins	no
<b>Court</b>	222	6	Atlantic	Congress	no
<b>Clinton</b>	394	3	Amity	Living.	no
<b>High. P.</b>					
<b>Kilmer</b>	884	10	Plainf.	Truman	no
<b>Raritan</b>	357	20	Third	Fifth	no
<b>Woodbr.</b>	409	20	Seventh	Eleventh	no

which represent a rectangular area defined by two latitude and longitude points. As soon as a vehicle enters the area defined by these points, the sensing vehicle starts collecting data and continues to collect while the vehicle is inside of the box. The trip boxes ensure that data is collected only from streets where it is important to estimate parking availability as opposed to areas where parking is available during the whole day. This data set was used in the [5] paper experiments.

### B. Brooklyn data set

The second data set was collected in downtown Brooklyn, NY, and it is used for the first time in this paper. Six sensing vehicles equipped with the ultrasonic sensor and Sony PS3 webcam were deployed to collect parking data during the fall of 2010, over four different work days. The starting location for each vehicle was NYU - Poly University, and each vehicle was assigned a different route to cover an area in downtown Brooklyn but some of the routes had overlapping areas. Figure 3b depicts the location of more than 1.5 million points collected during the experiment. Unlike the Highland Park data set, in Brooklyn the number of days in which vehicles were deployed was small but the size of the parking area was much larger. As a large, dense city, Brooklyn also represents a more complex parking landscape than Highland Park. In this data set, the concept of trip boxes was not used. From this large data set, we extracted eight streets in downtown Brooklyn for which we had at least three passes through the same street (Table I). In total, the number of data points collected by the [5] in both Highland Park and downtown Brooklyn areas and used in experiments was greater than 2 million.

## III. MAPPING ON-STREET PARKING SPACES USING AGGREGATION

Every time a sensing vehicle passes through a street, it can collect and report the rangefinder sensor measurement at each location. When several passes through the same street have been obtained, this will provide us with multiple snapshots of the parking occupancy in this street. To infer from this data whether a particular road-side spot is a legal parking space, our algorithm exploits the following key idea: *spaces that almost never have parked cars are probably not legal parking spots*

such as fire hydrants, private garages, side streets, or bus zones. On the other hand, *spaces which are frequently occupied almost every time are likely to be valid parking spaces*. Such information can be inferred by aggregating available time series. In the following sections we will describe our proposed weighted occupancy rate thresholding algorithm as well as two baseline algorithms. Before presenting these algorithms, we will discuss the pre- and postprocessing stages common to all these algorithms.

### A. Baselines, pre- and postprocessing

Since the sensor measurements from different passes are not necessarily taken at the exact same position, the first step in the preprocessing stage was to spatially discretize streets in one meter cells. Then, all obtained GPS readings and their corresponding sensor measurements are linked with the matching space cells. In the case that we have several sensor readings associated with one cell (for example a vehicle was standing on the certain spot for several minutes) we take the median value of the sensor readings from these locations. However, GPS readings are typically accurate up to three meters which can cause certain sensor readings to be associated with the wrong cells. In the next step we discretize sensor measurements. As described, the parking sensor provides measurements from 12 to 255 inches. Within this range it is possible to detect not only parked vehicles, but also other objects on the streets such as traffic lights, trees, trash cans, staircases, etc (Figure 4a). A common feature for all the objects is that they are usually positioned behind parked cars. As a first step, we weed



Fig. 3: Locations of the GPS traces for both data sets

out all unnecessary objects by applying a threshold method to time series; everything below the threshold is considered as a detected vehicle and it is assigned a value of 0 and everything above the threshold has a value of 1 and represents that nothing was detected on this space in the street. The threshold values vary for different streets; for streets without a bicycle lane this threshold is set as 100 inches and otherwise is set to 150 inches. Figure 4b demonstrates histograms of the detected vehicles for streets with and without bicycle lanes (easily obtained from Google Maps) which we used to select appropriate thresholds. Finally after the preprocessing step is finished we obtain several discretized time series on which we can apply aggregation algorithms.

As one of the baseline approaches we use the *trivial approach* which estimates that all parking spaces are valid similarly to [10]. The trivial approach can be useful to inform us about the percentage of illegal parking spaces in the street after it is compared with the ground truth parking maps.

Another *baseline approach* which we use to aggregate discretized time series is occupancy rate thresholding. It decides if a certain location is a legal/illegal parking space by taking the average of all time series for that specific location. If this calculated average is greater than 0.5, then this location is considered as a illegal parking spot and vice versa. The reasoning behind this approach is that if on average there were no parked cars on this location, there is a greater chance that this is not a legal parking spot. The reason why the threshold above is set to 0.5 is because even if a lot of places are no parking zones, people tend to temporarily park on fire hydrant or bus zones spots as is noticed in the data set.

At the end of the aggregation phase, the post processing step is applied to smooth out the resulting time series. The smoothing method finds all legal parking spaces in the resulting time series that are less than three meters and then converts them to illegal parking spaces. In the same fashion, it converts all illegal spaces to legal if they are less than three meters. The smoothing step is important because it eliminates all parking zones that are too small to be parking spots and vice versa. For example, the smallest non-parking zones are several garage entrances and side streets that are not smaller than 3 meters. With the post processing step small spikes that show up in the resulting time series are eliminated and the time series is smoothed. After smoothing we connect each space to its GPS

location and plot the map of legal/illegal parking spaces.

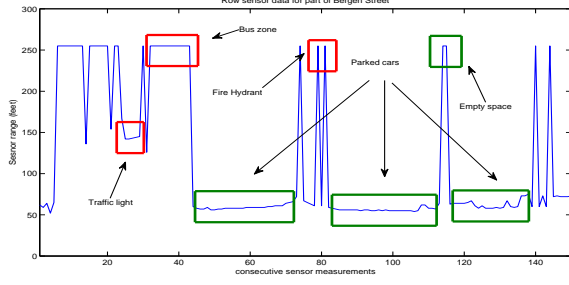
### B. Weighted occupancy rate thresholding approach

This algorithm is motivated by the observation that not all time series from the same street provide equally good information. The time series that have a larger occupancy of vehicles give us more valuable information as opposed to the time series which were collected while the streets were almost empty. The reasoning behind this is the following: when the majority of street parking spaces are occupied, only illegal parking spaces tend to remain empty. When many parking spots are available it is hard for our algorithm to distinguish spots that are empty because they are illegal from spots that are empty because of a lack of parking demand. The weighted occupancy rate thresholding approach utilizes this information to assign more weight to the time series that have high vehicle occupancy than to time series with low vehicle occupancy. An outline of the proposed algorithm is given in the Algorithm 1. The algorithm proceeds as follows: for a given input GPS and corresponding sensor traces the algorithm first performs preprocessing steps as described before. Then for every time series  $p$  we calculate the weight  $W$  as the ratio of the occupied slots to the total slots. In the next step we normalize the weights in order to distinguish between low and high informative time series. Then for every space cell  $i$ , the weighted average of sensor readings  $S$  for each time series  $p$  was calculated. In order to determine if the parking space is legal or not we apply the threshold method. Similarly to the baseline approach, the resulting aggregated time series is smoothed to eliminate small spikes. Finally all spaces are again linked to their GPS coordinates and can be plotted in order to obtain a map of legal/illegal parking spots.

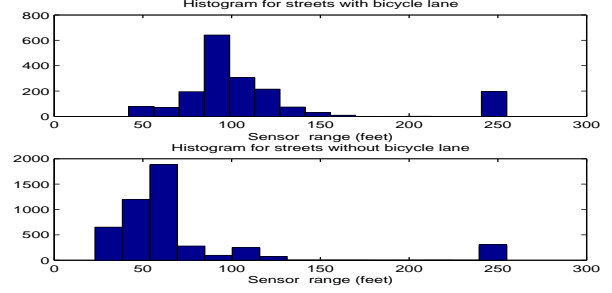
It is important to mention that, although this simple algorithm was developed specifically for the purpose of aggregating time series for parking map estimation, this algorithm can be generalized to other sensing systems which have multiple sensor readings (time series) from the same locations. For example, this approach can be used in aerosol retrieval [11], to estimate the level of aerosol in the air by combining satellite- and ground-based sensor measurements.

## IV. EXPERIMENTS

The baseline and proposed algorithm are evaluated on both the Brooklyn and Highland Park data sets. We compared



(a) Raw(before preprocessing) sensor data for the part of Bergen Street



(b) Histograms of the detected vehicles for streets with (top) and without bicycle lanes (bottom)

Fig. 4: Data properties

**input** : The GPS and sensor traces  $s$   
**output**: Array of legal/illegal parking cells

1. Discretize GPS traces into  $N$  equidistant cells
2. Assign each sensor value  $s$  to the corresponding cell  $i$
3. Apply threshold approach to obtain 0/1 sensor time series
4. Calculate weights for each time series  $p$

**for**  $p \leftarrow 1$  **to**  $P$  **do**

$$W_p = 1 - \frac{\sum_{i=1}^N s_{p,i}}{N};$$

**end**

5. Normalize the calculated weights

**for**  $p \leftarrow 1$  **to**  $P$  **do**

$$\tilde{W}_p = \frac{W_p}{\sum_{i=1}^P W_p};$$

**end**

6. Apply normalized weights to a each cell

**for**  $p \leftarrow 1$  **to**  $N$  **do**

$$\hat{s}_i = \sum_{i=1}^P \tilde{W}_p \cdot s_{p,i};$$

**end**

7. Apply threshold to estimate if cell is legal/illegal parking space

$$\hat{s}_i = \begin{cases} 1 & \text{if } (\hat{s}_i \geq \text{threshold}) \\ 0 & \text{otherwise} \end{cases}$$

8. Smooth out obtained time series.

**Algorithm 1:** The weighted occupancy rate thresholding algorithm

the output of both algorithms to ground truth parking maps. The ground truth maps are obtained by manually creating parking maps from satellite images and manually labeling areas as legal and illegal parking based on satellite and Google Street view imagery. In addition, ground truth maps are also discretized on the resolution of 1 m similarly to algorithm output maps. The dimensions of bus stops and entrances of parking garages are apparent from these images, making

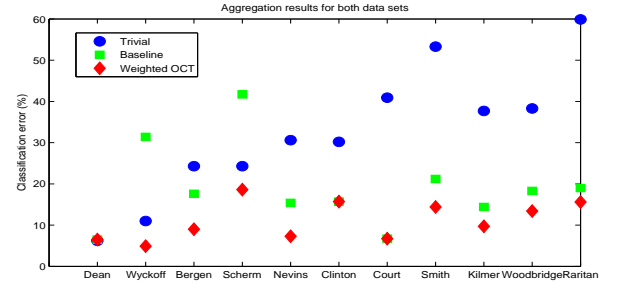


Fig. 5: Aggregation results for both data sets labeling straightforward. For fire hydrants, we labeled 5 meters (15 feet) on each side of the hydrant as illegal parking spots (according to NYC parking rules). Classification error was used as a measure of accuracy where the value of each predicted space cell is compared to the corresponding ground truth space cell. After evaluation we noticed that classification error is in the range of 5 to 15% depending on the street. The error is larger when we have a large number of illegal parking spaces such as fire hydrants, garage entrances, or bus zones that were hard to estimate. Another important factor which influenced the classification error is the number of diverse time series for specific streets; the more time series from different periods of the day we have, the easier it was to estimate if parking spaces are legal or illegal.

Figure 5 shows the results from eight streets in downtown Brooklyn. The weighted occupancy rate thresholding approach out-performs the baseline approaches for the majority of the streets. Furthermore, for some of the streets (ex. Wyckoff and Schermerhorn), baseline approach that averages all time series had very low accuracy (between 30 and 40%) which indicates that simple averaging of the time series is not a good approach. We observed that on several streets both algorithms have the same performance such as Dean, Court, and Clinton Streets. After inspection, we discovered that the time series for all passes for those streets were very similar (all taken on the same day over a close time period) which caused the weights for the weighted ORT approach to be very similar. In the case where all weights are similar, the proposed algorithm behaves the same as the baseline algorithm. When we compare street by street performance, we notice that the error varies significantly

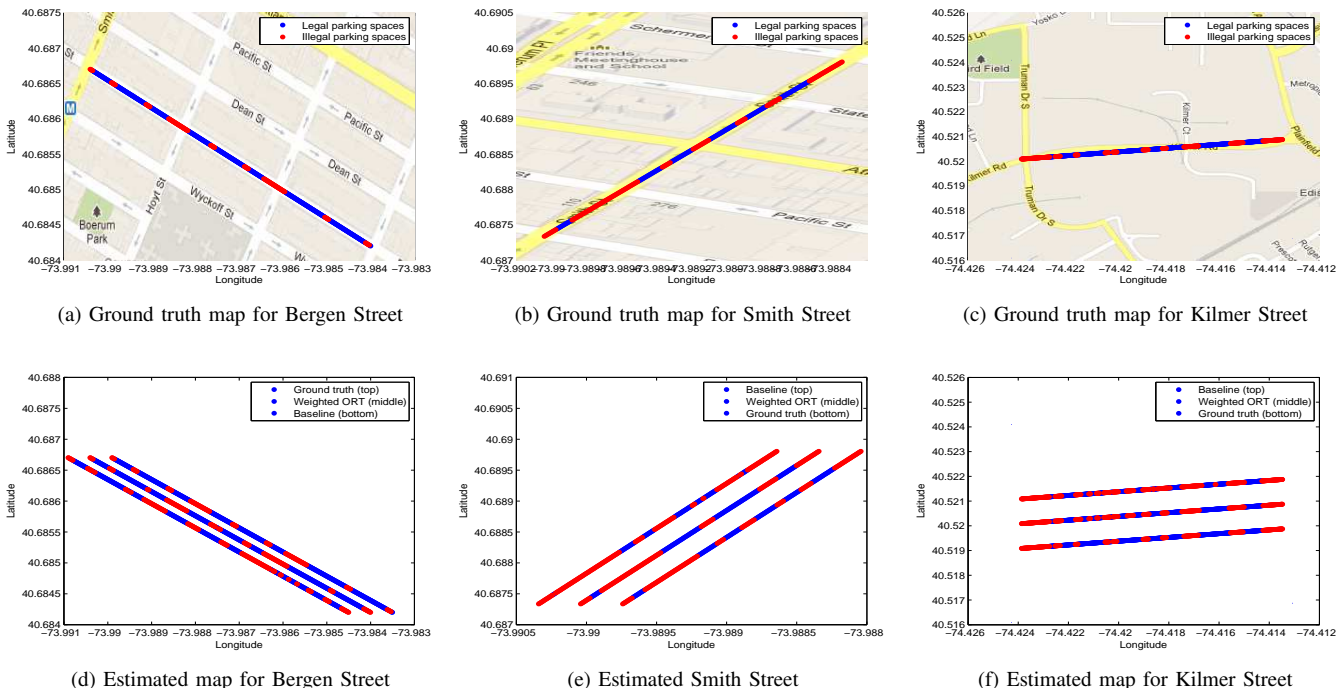


Fig. 6: Ground truth (top) and estimated (bottom) parking maps for three test streets

from street to street for both approaches due to significant distinction between the streets. For example, Dean Street is very narrow and the side where the sensor was measuring had no fire hydrants or bus stops, but only one garage entrance and three intersections. On the other hand, Schermerhorn Street has several fire hydrants, bus stops, and garage entrances. One of the reasons why the error was slightly larger on Smith Street is because parking rules were changed in the last couple of years and city authorities installed several new fire hydrants but left old parking markings visible. This caused many of drivers to park on illegal spots, making those spots hard to detect. Even during inspection of streets on Google maps we noticed quite a few parking violations (Figure 2b and Figure 2c).

Figure 5 also shows results for the Highland Park data set. Similarly to the Brooklyn data set, the proposed algorithm performs better than the baseline approach and trivial approach. Although we have a larger number of passes in Highland Park than in Brooklyn, estimation of the parking maps for Highland Park was more challenging since Highland Park streets are much less occupied and located in a low populated residential area where a lot of residents park on side streets or behind their houses. Thus, most of the time series contained very little information about parked vehicles, which made it difficult to distinguish between legal and illegal parking spaces after aggregation. Furthermore, Raritan Avenue has a large number of slotted parking spots which were rarely occupied, leading to a high percentage of non-parking spaces in that street which was hard to estimate. Finally, Figure 6 displays true and estimated parking maps for baseline and proposed

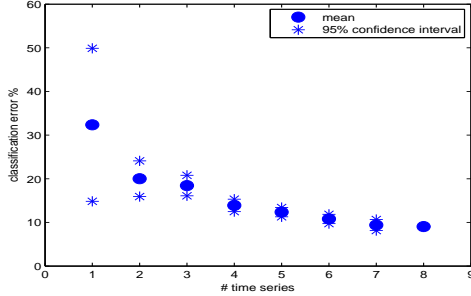
approach. From the figure we can observe that the weighted occupancy rate thresholding approach generates more accurate maps than the baseline approach and it is able to accurately detect most of the illegal parking spaces.

It will be interesting to see how the proposed algorithm compares with other approaches for parking map estimation, for example, using GPS traces in cars to see where they stop for long periods or perhaps participatory sensing where users mark spots as legal/illegal as they drive to find an open "spot" indicated by the system. However, we were not able to obtain data from these approaches in order to perform comparison.

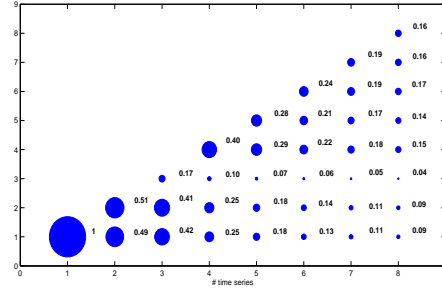
#### A. Properties of weighted occupancy rate thresholding algorithm

In the previous section we demonstrate that the proposed approach outperformed the trivial and baseline approaches. Now we investigate the behavior of the weighted occupancy rate thresholding algorithm when we change the number of time series used for aggregation. In the following experiment we used the time series for Bergen Street and we evaluated the performance of the proposed algorithm for each combination of time series. In Figure 7a, we can observe how both the classification error and confidence interval decrease when we add more and more time series. This indicates that the performance of weighted occupancy rate thresholding algorithm improves when we receive additional time series.

Next, we investigated how weights of the proposed algorithm change when we add more data. The weights for each time series in each iteration are presented in Figure 7b for one combination of Bergen Street time series. In the first iteration we have only one time series which has a weight of one. In



(a) Performance of weighted occupancy rate thresholding algorithm for different number of time series



(b) Weights behavior for different number of time series

Fig. 7: Proposed algorithm properties

the second iteration, we added additional time series which were similar to the first one and the algorithm estimated approximately similar weights of 0.49 and 0.51. As we add more time series the weights are decreasing, giving more weight to the more informative time series. Finally, after all time series were observed the algorithm assigns the final weight. It is interesting that time series number three received the smallest weight in all cases. After inspection, time series number three indicates that on that time of the day there were only a few parked vehicles on Bergen Street for that specific time. If only this time series is available it will be extremely hard to distinguish between illegal parking spaces and legal parking spaces that were not occupied for that time of the day.

Finally we examine the fraction of false positive and false negative for all test streets. The false negative rate in our case represents the number of illegal parking spaces that are classified as legal, and vice versa for the false positive ratio. Table II demonstrates these ratios for all test streets.

TABLE II: False negative and false positive rates

Street name	false negative (%)	false positive (%)
Bergen	3.29	5.76
Smith	8.36	6.11
Nevins	3.14	4.19
Dean	0.66	5.91
Scherm.	9.31	9.31
Wychoff	1.32	3.62
Court	4.50	2.25
Clinton	9.14	6.60
Kilmer	0.34	9.39
Raritan	11.7	3.92
Woodbr.	5.62	7.82
Average	5.21	5.89

Based on the Table II analysis of false negative rate, we notice that only  $\sim 5\%$  of parking spaces will be classified as legal even though they are illegal. This is very important for applications that suggest to users where they are allowed to park in the streets.

## V. DISCUSSION

In this section we will discuss in more detail some of the issues we encounter.

**Data sets limitations.** Due to the expensive data collection process we were able to evaluate the proposed approach only on dense urban areas where parking spots are typically full. In the case of small towns or more suburban areas where parking occupancy is very low, the proposed algorithm still needs to be evaluated. In the case of the Brooklyn data set, the main problem was that the data collection process lasted only a few days with small time intervals between passes through the same street. This resulted in all time series for a particular street holding very similar information. For example, for Clinton Street we had three passes of the same sensing vehicle in the time period of only a couple of hours. The parking situations did not significantly change and the algorithm concluded that all empty spaces at that time were illegal. This problem can be solved if more diverse data were available, especially in cases where most of the parking spaces were occupied (e.g. during night). On the other hand, in the Highland Park data set we had several time series during the day collected over a larger time period. The main difficulty in estimating illegal parking spots was the fact that the area is not as densely populated and parking spaces remain empty for most of the day, making estimation more challenging.

**Parking maps accuracy.** Based on results obtained from the evaluation we can pose the question of whether  $\sim 90\%$  accuracy is sufficient to use this technique in practice. For applications such as detection of zones with large numbers of illegal spaces for travelers or DOT departments, an accuracy of  $\sim 90\%$  is more than sufficient. On the other hand, for GPS receiver or PARKNET applications which require fine resolution of parking maps, accuracy of the proposed approach is adequate, especially if we take into account that on average only  $\sim 5\%$  of the error is due to false negative rate. This implies that situations where illegal parking space is estimated as legal and which may cause user frustration will be very rare.

**Multilane roads.** An important issue we encounter is how to detect legal/illegal parking spaces when sensing vehicles are passing through multilane streets. Unfortunately, current GPS receivers are not precise enough to distinguish in which lane the vehicle is driving. In the data set used in the experiments, most of the streets are single lane streets and in cases where we

have multiple lane streets we eliminate the time series in which the sensing vehicle is changing the lane. In the process of data analysis, we noticed that if the vehicle is not in the right lane, sensor measurements are unusually low (the sensing vehicle is detecting another vehicle that is very close) or unusually high for longer time periods (the sensing vehicle did not detect anything because it was far away). Developing an approach which will automatically detect these situations remains for future work.

**Privacy.** When we deal with GPS trace data sets, one of the main issues is the question of privacy. Sensing vehicles must reveal their position to the parking estimation system. This can lead to the users home and work locations being revealed simultaneously with their daily patterns. The proposed approach [12] addresses the problem of preserving privacy for crowdsensing and it can be applied to our problem. In this paper we do not further address this issue.

## VI. RELATED WORK

In recent years several systems have been developed and tested for parking space monitoring. Parking garages use systems which count the number of vehicles entering/exiting the garages and display an estimated number of empty parking spaces on the garage entrance message signs [13]. A couple of interesting approaches were recently proposed, where users can buy and sell privately owned parking spaces [14]. Recently two systems which monitor on-street parking spaces were proposed. The first one is SF-park [4], a project in the city of San Francisco which employs a large number of fixed sensors for parking space detection and second, the PARKNET project [5], where low cost ultrasonic sensors coupled with a GPS receiver were installed in sensing vehicles in order to detect on-street parked vehicles. It is assumed that parking maps are already available to the system as noted in the introduction. The construction road maps from GPS traces are recognized as an important problem and in the past couple of years several studies have been done on this topic. Several data mining algorithms such as K-means, kernel density estimation or trace merging algorithms [7], [15] were proposed in order to estimate road maps from low resolution and low sampling traces. As opposed to these approaches, the community in Open-StreetMaps (OSM) [16] is manually extracting roads from arterial images and GPS traces. However, all proposed approaches focus only on estimating road maps, leaving the problem of creating parking maps open. There is a wide variety of applications for traffic state estimation that use crowdsensing data such as the detection and prediction of traffic light schedules using smart phone images [17], providing real time bus arrival times by distinguishing buses using microphone and GPS sensors [18] or developing fuel efficient maps using fuel consumption sensor data [19].

## VII. CONCLUSIONS

In this paper we have presented how crowdsensing can be used in the mapping of on-street parking spaces to construct legal/illegal parking maps. Based on the more than 2 million

points of parking sensor readings in our experiments we reached the following conclusions: First, we show how the parking maps can be estimated with an accuracy of  $\sim 90\%$  from the parking sensor data using proposed weighted occupancy rate thresholding algorithm. Then, we demonstrated that the proposed approach outperforms trivial and baseline approaches on both data sets. Finally, we illustrate how accuracy of the proposed algorithms improves by adding more passes (time series) through the same street.

## ACKNOWLEDGMENT

The authors would like to thank Patrick Yuen for his help in data preprocessing and Jelena Coric for her help in editing the paper.

## REFERENCES

- [1] Intelligent transportation systems. [Online]. Available: [www.its.dot.gov/](http://www.its.dot.gov/)
- [2] D. Shoup, "Cruising for parking," *Transport Policy*, vol. 13, no. 6, pp. 479–486, 2006.
- [3] —, "The price of parking on great streets," University of California Transportation Center, Tech. Rep., 2011.
- [4] Sf-park project. [Online]. Available: <http://sfpark.org/>
- [5] S. Mathur, T. Jin, N. Kasturirangan, J. Chandrasekaran, W. Xue, M. Gruteser, and W. Trappe, "Parknet: drive-by sensing of road-side parking statistics," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 2010, pp. 123–136.
- [6] Google maps. [Online]. Available: <https://maps.google.com/>
- [7] J. Biagioni and J. Eriksson, "Inferring road maps from gps traces: Survey and comparative evaluation," in *Transportation Research Board 91st Annual Meeting*, no. 12-3438, 2012.
- [8] X. Liu, J. Biagioni, J. Eriksson, Y. Wang, G. Forman, and Y. Zhu, "Mining large-scale, sparse gps traces for map inference: comparison of approaches," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 669–677.
- [9] G. Brosicke, O. Mayer, R. Eri, and H. Seeger, "The automatic parking brake," *ATZ Automobiltechnische Zeitschrift*, vol. 103, pp. 39–42, 2001.
- [10] Primo spot. [Online]. Available: <http://www.primospot.com/>
- [11] K. Ristovski, S. Vucetic, and Z. Obradovic, "Uncertainty analysis of neural-network-based aerosol retrieval," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 2, pp. 409–414, 2012.
- [12] K. Tang, P. Keyani, J. Fogarty, and J. Hong, "Putting people in their place: an anonymous and privacy-sensitive approach to collecting sensed data in location-based applications," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 93–102.
- [13] Smart-parking at rockridge bart station. [Online]. Available: <http://www.path.berkeley.edu/path/research/featured/120804/smart-park.html>
- [14] Mobileparking. [Online]. Available: <https://www.mobileparking.com>
- [15] L. Cao and J. Krumm, "From gps traces to a routable road map," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2009, pp. 3–12.
- [16] Open streetmap. [Online]. Available: <http://www.openstreetmap.org>
- [17] E. Koukoumidis, L. Peh, and M. Martonosi, "Signalguru: leveraging mobile phones for collaborative traffic signal schedule advisory," in *Proceedings of the 9th international conference on Mobile systems, applications, and services*. ACM, 2011, pp. 127–140.
- [18] P. Zhou, Y. Zheng, and M. Li, "How long to wait?: Predicting bus arrival time with mobile phone based participatory sensing," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 379–392.
- [19] R. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. Abdelzaher, "Greengps: A participatory sensing fuel-efficient maps application," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 2010, pp. 151–164.