Addressing Parking Challenges in Downtown Pittsburgh

Tayo Fabusuyi and Robert Hampshire

Abstract

This paper discusses the development of ParkPGH, a novel smart parking application that provides real time and predictive information on garage parking availability in downtown Pittsburgh. The initiative is in response to the increased demand for parking spaces in downtown Pittsburgh and the desire to improve drivers’ parking experiences. The application includes a predictive model that uses as input historical parking, weather and event data to provide estimates of available parking spaces. We provide an example of the model implementation using data from the Theater Square garage where we utilize neural network-based predictors and multiple net searches to generate estimates of parking availability. Provision was also made for binary classifiers given the need to reduce the possibility of Type II errors. Outcome measures show that more than 50% of respondents reported a reduction in parking search time with the magnitude ranging from a minute to more than six minutes.

T. Fabusuyi (Corresponding author)
Numeritics, Pittsburgh, PA and Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA.
Email: tfabusuyi@cmu.edu

R. Hampshire
Transportation Research Institute (TRI), University of Michigan, Ann-Arbor, MI.
Email: hamp@umich.edu
1. Introduction

The past decade has witnessed the resurgence of interest in cities. US Census data, show that Americans are increasingly choosing to move to cities with the wave in urban growth concentrated in and around city centers especially in the West and the South (Frey, 2010). This trend in increased urbanization is not specific to the United States (US). The United Nations Population Fund (UNFPA) observed that the year 2008 is the first time in history when more than half of all humans will be living in urban areas. In absolute terms, this number is expected to rise to 5 billion by 2030. It is especially noteworthy that within a span of 50 years, there has been a 20% increase in the global urban population, from 34% in 1960 to an estimated 54% in 2014 (United Nations Department of Economic and Social Affairs, Population Division, 2014), (United Nations Population Fund (UNFPA), 2007).

However, as a result of weak finances, an appreciable number of cities have limped along for decades, delaying much needed investments in critical infrastructure. Urban infrastructure deficits are further exacerbated in cities’ central business districts (CBDs) given the increased concentration of workers and visitors. The City of Pittsburgh provides a canonical example in this regard. Less than 25% of the 288,000 individuals who work in Pittsburgh reside within the city. In addition, the compact, 0.5 square mile area CBD is becoming increasingly attractive as a residential area and continues to be a mecca for cultural events and sporting activities (Fabusuyi & Hampshire, 2013). Similar issues are being witnessed in an appreciable number of cities across the US.

The present work addresses the associated parking problems with an emphasis on Pittsburgh’s downtown. We developed ParkPGH, a smart parking information system that uses parking, event and weather data to provide information on the availability of parking within downtown Pittsburgh with the goal of reducing parking search time and search time variability. The development of the product includes a robust needs assessment, an open source platform, a detailed evaluation component, the use of a modular design, and a predictive algorithm.

The balance of the paper discusses each of these features. Section 2, assesses the nature and extent of Pittsburgh's parking problem through a stakeholder analysis and a broader environmental scan. Insights from the needs assessment are used to inform Section 3, the systems development section. The systems development section identifies the intervention that we determined would be the most robust in tackling the deficit identified in the needs assessment section. Section 4, the decision analytics section, docu-
ments the prediction approach and presents findings from the analysis. Section 5, the evaluation section, focuses on post-deployment evaluation, and itemizes the management challenges encountered during the project’s implementation. It also presents the evaluation framework, the data sources employed, and estimates of the project’s impact using key outcome metrics and cost effectiveness measures. Section 6, the conclusion section, summarizes the work, provides insights on the value-added by the application and provides suggestions for further research.

2. Needs Assessment

Parking space is at a premium in downtown Pittsburgh, an area of 0.5 square miles with a workforce strength of approximately 130,000 (Pittsburgh Downtown Partnership, 2012). Apart from the geographical limitations, a number of additional factors explain this situation. As a result of policy measures put in place during the 1990s that were motivated in part by the Pittsburgh Downtown Plan (Strada, 2009), there has been a noticeable decrease in the supply of available parking spaces over the last two decades. In addition, current and proposed developments are anticipated to further reduce the total available number of parking spaces. These developments, and the need to avoid a situation where parking becomes a binding constraint to the economic vitality and growth of the downtown area, prompted both the Pittsburgh Cultural Trust (PCT) and the Pittsburgh Downtown Partnership (PDP) to seek solutions to these parking problems.

Addressing the parking problem demands a robust approach that can establish the nature of the problem, design and implement a program intervention to rectify the problem and provide an assessment of the degree to which the problem has been ameliorated. The needs assessment speaks to the first of these requirements. Our approach to the needs assessment involves a gap analysis where the difference between the desired state and the present state determines the need. Stakeholders’ expectations are examined for feasibility by reflecting the broader environmental constraints in the analysis. The former is made up solely of primary data while the environmental scan was conducted using secondary data from US Census and document review.

Key stakeholders, including the PCT and PDP, were identified based on their roles, the resources they control or the responsibilities they have and the relationship they have within the broader transportation ecosystem. In addition to the detailed conversations with key stakeholders, intercept and online surveys were administered to 736 individuals in order to have an assessment of individuals’ parking experiences and to gather baseline data.
on key indicators. Table 1 provides the summary results for the baseline survey data on the key objectives identified by stakeholders.

Table 1: Baseline Data on Key Program Objectives

<table>
<thead>
<tr>
<th>Program Objectives</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parking search time/Search time variability¹</td>
<td>7.3min</td>
</tr>
<tr>
<td>Late coming incidence</td>
<td>27.0%</td>
</tr>
<tr>
<td>Perception about parking (% indicates those surveyed without a positive response)</td>
<td></td>
</tr>
<tr>
<td>Parking satisfaction</td>
<td>25.7%</td>
</tr>
<tr>
<td>Ease of finding a parking space</td>
<td>22.4%</td>
</tr>
<tr>
<td>Overall parking experience</td>
<td>22.7%</td>
</tr>
</tbody>
</table>

From the combined primary data, we were able to obtain key insights as to the state of parking within downtown Pittsburgh. We found that the problem is less an absolute shortage of parking spaces but more of a limited parking availability in close proximity to individuals' destinations. Secondly, there were numerous complaints about low turnover in parking spaces – a situation attributed to employees parking in prime parking spots. In addition, the need to emphasize off-street parking became clear given that within the downtown area only 13% of all available parking is on-street. Finally, there was a consensus on the need to focus on commuters, a viewpoint that kept resurfacing throughout the semi-structured interview sessions.

![Figures 1(a) and (b): Mean and variance, respectively, of the number of Vacancies on weekdays, weekends and holidays. Data obtained from the gate counts of Theater Square garage over a 600 day period.](image)

¹ Parking search time is measured by the mean search time. This measure is also used as a proxy for search time variability or the dispersion around the mean given that the search time data has an exponential distribution.
In light of these findings, the parking and travel demand pattern of commuters merits a careful look. Figure 1, obtained from a representative parking garage, illustrates parking demand using the average available parking spaces for weekdays and weekend-holidays and their corresponding variances. The huge drop in the number of available spaces observed between 10 am and 3 pm on weekdays is considered to be work-related given that the number of spaces is relatively stable – i.e. low variance. In contrast, we consider the drops around 3 pm and close to 8 pm on weekends to be event-driven because the number of parking spaces fluctuates greatly (i.e. high variance) depending on event occurrences. These insights, coupled with findings from the needs assessment phase were subsequently used in determining the robust program intervention. They provided the rationale for framing the research question – can a demand-side intervention, specifically providing real-time and predictive information address the (perceived) lack of parking availability in downtown Pittsburgh? The rest of the paper, starting with the schema in Figure 2, is geared towards answering this question.

Figure 2: Smart Parking Application Schema
3. Systems Development

The program intervention is a pilot product called “ParkPGH,” a smart parking application providing real-time and predictive information on the availability of off-street parking within Downtown Pittsburgh. The app falls within the realm of initiatives classified into two broad sections: parking guidance systems, and real-time/prediction information. Parking guidance systems (PGS) use variable message signs (VMS) to inform drivers about available parking spaces with the information being relayed through a series of channels and oftentimes, integrated with traffic information system. Much of the literature on parking guidance systems is concerned with transit and park-and-ride lots (Shaheen & Kemmerer, 2008) with examples of the system described in Orski (Orski, 2003). Another stream of work on PGS explores their use inside of parking facilities (Caicedo F.). ParkPGH is distinct from the parking guidance system literature in two ways: ParkPGH is not coupled with transit, and it does not employ VMS. The parking availability information is available only through mobile devices, interactive voice response (IVR) and the Internet.

Real-time and prediction information systems examine the display and use of information for finding parking spots. Information on parking availability is either provided during a trip or before the trip begins. The works of Caliskan et al. (Caliskan, Barthels, Scheuermann, & Mauve, April 2007) and Teng et al. (Teng, Qi, & Martinelli, 2008) are examples of systems that provide parking prediction models based on information exchanged between wirelessly connected vehicles for use during a trip. Other variations include the use of agent-based modeling in simulating drivers’ behavior (Martens & Benenson, 2008); an allocation model by Teodorovic and Lucic (Teodorovic & Lucic, 2006) that accepts parking requests via an inventory control system; utilizing sensors to provide real time parking data (Vlahogianni, Kepaptsoglou, Tsetsos, & Karlaftis, 2014) and predicting the number of available parking spaces where the parking requests are routed to a number of competing parking facilities (Caicedo, Blazquez, & Miranda, 2012).

Our approach is to employ a predictive model that utilizes historical garage occupancy rates with detailed information on events and weather conditions. The approach improves on similar works that have been bedeviled with poor forecasts or the inability to determine the ideal number of neural units (Yang, Liu, & Wang, 2003). The prediction model utilized for the present study is an event-based parking prediction model for use before a trip begins along with historical parking and event data to predict future parking availability. These predictions have been shown to reduce the uncertainty.
often related to parking in downtown areas and central business districts (Bos, Ettema, & Molin, 2004).

The application was implemented by combining systems development and integration with a parking prediction algorithm as shown in Figure 2. The system development and integration module collects real-time parking information from both public and privately held parking garages. This was made possible through the use of a web application programming interface (API) and infrastructure that collects, validates, and stores parking information in real time. The system integration also includes the development of an iPhone application, text message gateway, and an API that provides third party developers access to ParkPGH data.

The prediction model uses as inputs historical parking and event data that have occurred downtown and provides estimates of the available parking spaces for each garage. The prediction model is trained on a historical parking data set. This dual-pronged technological innovation was deployed through a pilot program that monitors eight parking garages totaling 5000 parking spaces, representing about 20% of the total parking supply in downtown Pittsburgh. Parking information is updated every minute and delivered through multiple channels that include websites, iPhone app, SMS text, voice and a mobile version of the website that provides the same information.

![Figure 3: Pittsburgh’s downtown map showing available parking spaces](image-url)

Figure 3: Pittsburgh’s downtown map showing available parking spaces
as the traditional website but is optimized for mobile devices such as Blackberries and Android phones. We have embraced the traffic sign colors in providing information to patrons looking for parking spaces. The green, yellow, and red color-coding is complemented with a numerical figure that shows the available number of parking spots, except in cases where the garage is deemed full or close to full capacity. A snapshot of the website showing destinations within the Cultural District, garages and the available spaces is provided in Figure 3. The exact number of available parking spots is not shown when the garage is deemed nearly full.

Figure 4 is a screenshot of the ParkPGH iPhone application. In the pictured scenario a popular garage, Theater Square, is currently designated as “Near Full.” In addition to this real-time information, a plot of predicted parking demand is provided on the lower half of the screen. The predicted parking demand plot shows the average or baseline parking demand for the garage based on historical data. Additionally, the demand exceeding the average is also provided. In this scenario, the excess demand is predicted based on two events occurring near Theater Square garage that influence future parking availability.

4. Decision Analytics

The prediction models are presented in two parts. First, we focus on predicting the number of available parking spaces at a given time, based on a set of events and weather data using neural network-based predictors.

---

2 Sections 3 draws from a journal article by (Fabusuyi T., Hampshire, Hill, & Sasanuma, 2014) from Interfaces, the practice journal of the Institute for Operations Research and Management Science (INFORMS).
Given that the driver would like to know whether the garage is full, we provide a robust approach that reduces the possibility of Type II errors—a situation that occurs when the garage is full, but the application shows that parking spots are available. This motivates the classification methods presented in the second part of this section. A range of classification and prediction methods, including logistic regression, naïve Bayes classifiers, classification and regression trees (CART), and a neural network, complements the continuous prediction methods.

4.1 Data Description

Here we describe the data utilized for the predictive model using parking availability data for Theater Square garage, the analysis and dataset for other garages are similar. The training set of parking data includes the number of available parking spaces for each 10-minute interval for 18 hours each day from November 9, 2008 to July 10, 2010 as the dependent variable. The prediction model estimates parking vacancy based on the variables itemized in Table 2 below. Events are categorized as (1) morning (before noon), (2) day (12:10 pm to 4 pm), or (3) night (after 4:10 pm).

<table>
<thead>
<tr>
<th>Categorical independent</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theater Dummies: ben2, ben3, byh2, byh3, or2, or3, hnz2, hnz3</td>
<td>Dummies for theater events at the Benedum Center (ben*); Byham Theater (byh*); O’Reilly Theater (or*); and Heinz Hall (hnz*)</td>
</tr>
<tr>
<td>Sport Dummies: pir2, pir3, hnzf2, hnzf3, pen2, pen3, stl2, stl3</td>
<td>Dummies for sporting events - Pirates (pir*); Heinz Field (hnzf*); Penguins (pen*); and Steelers (stl*)</td>
</tr>
<tr>
<td>Day of the week: Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, holiday</td>
<td>Dummies for day of the week</td>
</tr>
<tr>
<td>Time of the day: period</td>
<td>Period of the day measured in 10-minute increments</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numeric independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather: snow, rain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numeric dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
</tr>
</tbody>
</table>
4.2 Neural Network and Measures of Predictive Accuracy

During preliminary analysis, a multiple linear regression analysis proved to have low predictive power. Thus, we sought better prediction models using neural networks. The forms of neural network approaches employed include both the generalized regression neural network (GRNN) (Specht, 1991) and the multilayer feed-forward network (MLFN) (Svozil, 1997). The best net search from the architectures utilized is the GRNN numeric predictor. We carried out 31 trials for the training set. Measures of how close the predicted values are to the eventual outcomes are provided in the form of root mean square error (RMSE), mean absolute error (MAE), and standard deviation of the absolute error values for both training and testing trials. In addition, we provide measures for the percentage of bad predictions. A bad prediction value of 11.4% was obtained for the testing case. A prediction is deemed a bad prediction if it is more than ±5% of the absolute value of the actual parking availability.

Multiple nets were trained with predicted accuracy measures using RMSE ranging from 144 for the 2 Node MLFN to 60 for the GRNN. We also carried out sensitivity analyses to determine the reliability of the predictions for each testing case and to avoid over-fitting the training dataset. Results from the sensitivity testing were invaluable in estimating reliability measures as a result of changing the size of the subset of data used for testing and in ascertaining the quality of the predicted values. The ideal percentage testing case was found to be 20% with the RMSE value ranging from a low of 60.78 to a high of 62.90 for this threshold.

Figure 5 provides the rationale for the use of classifiers. The $45^\circ$ line represents the locus of points where the predicted and the actual values are of the same magnitude. The figure shows that the GRNN net provides a good fit in situations where the utilization of the garage parking spots is neither low nor high. However, the magnitude of the errors seem to increase at the extremes. The predicted values systematically overestimated

![Figure 5: Scatterplot comparing predicted values to actual parking availability](image-url)
the actual parking availability for low values and consistently underestimated the actual number of available parking spaces for high values. This explains our rationale for switching to a categorical dependent variable, especially at high-capacity utilization where users of the application may be extremely sensitive to Type II errors—a false negative (not full) when it is indeed full. To be conservative, we selected a threshold level of 85% for dichotomizing the dependent variable.

4.3 Predictive Classification Methods

As a complement to the continuous variable prediction models, we explored classification methods based on machine learning using a binary dependent variable—full or not full. We employ the same data set for the continuous availability prediction, which contains 36,949 observations of not full, representing 94 percent of the observations, and 2,409 observations of full. The independent variables remain the same as in the continuous variable prediction models. We report the prediction results of three classifiers, naïve Bayes classifier (NBC), logistic regression and classification and regression tree (CART).

Both the NBC and the Logistic Regression predict the class, (full or not full), using an approach similar to an empirically driven maximum-likelihood estimator (Mitchell, 2010). However, in contrast to these approaches, the classification and regression tree (CART) method is a nonparametric method, which employs a binary tree for classification and prediction (Loh, 2011). The classification methods were implemented using tenfold cross validation. Table 3 shows the results—average RMSE, precision, and recall. We note that the RMSE values obtained for the classification-based prediction methods are of much smaller magnitude compared to the numeric predictor because of the binary nature of the dependent variable.

Apart from the RMSE values, precision and recall measures are also provided. As shown in Table 3, the performance of CART is superior to the other approaches. This is not surprising given the independence assumption of both the naïve Bayes and the logistic regression. CART does not make this assumption and is free to build trees to exploit any correlation in the feature structure. The resulting CART has 341 nodes with 171 leaf nodes. Each leaf node corresponds to a unique combination of features or scenarios.
The CART extension to the GRNN numeric predictor provides a robust prediction platform. The hybrid approach mimics the real-time feed that is currently running across multiple channels of the smart-parking application in which the “full” sign is displayed when the parking space utilization goes above a specific level. The frequency of updates to be made to the models is determined by the levels of RMSE, percentage of bad predictions, precision, and recall values associated with the models. These thresholds will be established using weekly live predictions and analyzing the residual values. Input will also be solicited from garage operators regarding the tolerable level of error.

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>Precision</th>
<th>Recall</th>
<th>Binary outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve Bayes</td>
<td>0.252</td>
<td>0.951</td>
<td>0.964</td>
<td>Not full</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.299</td>
<td>0.237</td>
<td>Full</td>
</tr>
<tr>
<td>CART</td>
<td>0.129</td>
<td>0.984</td>
<td>0.991</td>
<td>Not full</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.842</td>
<td>0.756</td>
<td>Full</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.222</td>
<td>0.942</td>
<td>0.995</td>
<td>Not full</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.449</td>
<td>0.059</td>
<td>Full</td>
</tr>
</tbody>
</table>

The CART extension to the GRNN numeric predictor provides a robust prediction platform. The hybrid approach mimics the real-time feed that is currently running across multiple channels of the smart-parking application in which the “full” sign is displayed when the parking space utilization goes above a specific level. The frequency of updates to be made to the models is determined by the levels of RMSE, percentage of bad predictions, precision, and recall values associated with the models. These thresholds will be established using weekly live predictions and analyzing the residual values. Input will also be solicited from garage operators regarding the tolerable level of error.

5. Pilot Program Evaluation

In order to improve upon and ascertain the value added by ParkPGH, an evaluation of the pilot program was undertaken. Count data was used to track output measures that include the weekly usage volume for each of the delivery channels used to provide information by ParkPGH. This includes iPhone app, mobile and traditional website usage, number of text messages sent on request, bounce rate, average duration of page views, number of unique views and number of automated phone responses. Findings from the evaluation revealed that approximately one out of every two respondents reported that the application has reduced the time it takes them to find a parking space. The magnitude of the reduction in search time ranges from
as little as a minute to more than 6 minutes with the majority of individuals reporting a 4-6 minute reduction in search time.

In addition, process measures were utilized for formative evaluation purposes. Information obtained from these measures was used to make modifications to the smart parking project. Ease of use, difficulties with design and accuracy of the information provided are some of the process related measures tracked. A negative response on the online survey to any of these measures prompts an open-ended question that allowed the respondent to provide detailed information as to the nature of the problem being encountered. Such information was subsequently relayed to the development team.

However we are not able to report on the reduction in search time variability and patron perception post-deployment due to lack of data. The inability to provide outcome measures on some variables is not the only shortcoming associated with the data paucity issue. For example, it was not possible to establish conclusively whether there was response-shift bias in the self-reported reports. Finally, it is important to note that the figures obtained for the post-deployment measures were obtained from a fairly small sample size that may not be representative of the population of interest.

5. Conclusion

This research introduces an innovative demand-side intervention to address parking challenges in downtown Pittsburgh. The intervention, a smart parking application, uses garage occupancy rates with detailed information on events and weather conditions to provide real-time and predictive parking information. The pilot program has been effective in reducing parking search time and in changing the perceptions of patrons about the downtown parking situation.

Compared to other smart parking applications, ParkPGH is distinguished by a number of unique and distinct features. For one, a detailed stakeholders’ analysis was carried out, which was complemented by a process and summative evaluation. This served multiple purposes; it allowed stakeholders to identify shortcomings in the parking situation within downtown Pittsburgh, and determined the most robust form of program intervention to address the deficits. This comprehensive stakeholders’ analysis and the input from end users helped counteract the typically “silod” approach by which parking data is managed, an approach that precludes the exploration of synergies across different platforms and often leads to sub-optimal outcomes. In addition, involving stakeholders in the program design phase and having a robust evaluation platform yielded crucial insights on the design of product features.
and allowed modifications to be made to these features in real time. These provisions were invaluable in making the application user-friendly and effective.

The application’s relatively low cost, its ease of retrofitting, modular structure and its open-source platform can enable other cities to lower the costs of implementing and managing similar smart-parking solutions; significantly shorten their learning curves and ensures that transitioning to a fully-fledged integrated parking application could be achieved at a relatively low cost. The design process placed emphasis on creating citywide solutions that span both public and private providers and made the case for a coordinated approach that ensures that the application provides standardized, platform-independent solutions that could be scaled up – an invaluable feature that could promote its buy-in across multiple cities. Apart from the aforementioned, the modular design of ParkPGH makes provision for product enhancements, ensuring that retrofitting can occur with relative ease.

The novelty of the application is the deployment of the first real-time and predictive parking analytics system with input from multiple stakeholders. While the prediction method used in ParkPGH is not new, the environment in which it is deployed is unique and the approaches used in addressing these challenges have been nothing short of innovative. More importantly, of all the components of the smart parking application, the predictive module holds the most promise given its potential influence on commuters’ travel and parking demand patterns. Integrating parking information with traffic flow patterns is more effective if the parking information provided is not only real time but also predictive. Having this knowledge can encourage commuters to change the way they travel or how they schedule their trips. Apart from providing information on the demand side, garage operators could use the predictive information to better manage their facilities. These findings underscore the need to prioritize the implementation of a fully functional predictive module that encompasses both long- and short-term predictive modules.

References


