Under 2nd Round at Interfaces

Abstract

ParkPGH is a novel parking application that provides real-time and predictive information on the availability of garage parking spots within Pittsburgh's Cultural District. The core of the application is a systems development and integration module that collects real-time parking information from the garages by tapping into their gate counts. The real-time component is complemented by a predictive module that uses historical data and an events calendar to make predictions about parking availability. Visitors to downtown Pittsburgh have found the application useful in finding parking spaces; in 2011, drivers used ParkPGH more than 300,000 times to decide when and where to park. The application has also been beneficial to garage operators through the information it provides on parking demand that affords operators greater flexibility in addressing contingencies and managing lease holders.

The deployment of ParkPGH, which includes a robust evaluation component, is one piece of a broader transportation ecosystem within the Greater Pittsburgh region. Lessons learned from the initiative, along with the application's relative low cost, ease of retrofitting, and its open source platform will allow other cities and metropolis to significantly shorten their learning curves and lower the costs of implementing and managing similar smart parking solutions.

Keywords: smart parking, public sector, analytics, informatics, queueing theory, multi-stakeholders

Introduction

ParkPGH is a smart parking project designed to give real-time information on the availability of parking to patrons within the Pittsburgh Cultural District, a geographically-designated area within the city of Pittsburgh. Since its inception, the PCT has witnessed significant increases in attendance and patronage within the Cultural District. Attendance at performances within the District surged from 570,000 in 1990 to 2 million people in 2008, an increase of more than 250%. This development has placed considerable strain on the existing amenities within the District, particularly parking facilities, a situation further compounded by the scale of activities on the North Shore with their added demand for parking from sporting fans.

A series of initiatives have been put in place by the PCT over the last few years to help alleviate this problem, primarily on the demand side but very little value added was observed from these interventions. These initiatives included promoting fringe parking on the North Shore and using shuttle buses to transport patrons to event venues. This concept was discontinued because very few patrons made use of the facility. The Trust also offers pre-paid parking to its high-level patrons although the pre-paid parking is not bound by contract and the provider of the service is not compelled to provide parking spots. As a result, the PCT is reluctant to widely publicize this service.

To address this problem, PCT, with generous funding from the Benter Foundation, initiated ParkPGH, www.parkpgh.org, a smart parking, technology based program within downtown Pittsburgh. The program, which went live in December 2011, enhances the existing off street parking facilities within the District by providing real-time information using a host of information delivery methods that includes an iPhone application, traditional and mobile website, text messaging and an interactive voice response system. The primary goals of the program are to reduce search time and search time variability when finding a parking space within the Cultural District and to make the District a more desirable destination for patrons by reducing the anxiety and uncertainties related to parking issues. A secondary goal is to attract new patrons who were previously deterred by the uncertainty of parking availability. The pilot program currently monitors eight parking garages totaling 5000 parking spaces, representing approximately 20% of the total parking supply in downtown Pittsburgh and over 90% of the parking supply in the cultural district. This relatively large market share of parking spaces in the Cultural District provides a unique opportunity to evaluate the impact of a smart parking system.

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 approach used in ParkPGH is not new, the environment within which it is deployed is unique. The authors are unaware of any parking application in existence that has an operational predictive module. Additionally, involving key stakeholders in the program design phase and having a robust evaluation platform have yielded crucial insights on the design of product features and have allowed modifications to be made to these features in real-time. These provisions have been invaluable in making the application user-friendly and effective.

The rest of the paper is organized as follows. The next section provides a rationale for and presents the background context within which the project was conceptualized. The third section discusses the stakeholders' analysis component of the study. Section 4 addresses the systems architecture, documents the prediction approach and and itemizes the management challenges encountered during the project's implementation. The fifth section presents findings from the evaluation. The concluding section recaps and provides insight both on the approach and the value added by the application.

Background Context

Pittsburgh's downtown, nicknamed the Golden Triangle, is a geographical area of approximately 0.5 square miles area bounded by Grant Street and both the Allegheny and Monongahela rivers. The rivers act as natural boundaries and limit the area within the triangle. The area serves as the core of the business and economic activity for South Western Pennsylvania and its workforce strength of approximately 130,000 represents 32% of the City of Pittsburgh working population (Partnership, 2012). Available parking options within the triangle include on-street metered parking and off-street parking made up primarily of garages and surface lots. The garages are operated or owned primarily by two entities - the Pittsburgh Parking Authority that is responsible for most publicly owned garages and Alco Parking, the largest operator of the private parking garages.

As a result of policy measures put in place during the 1990s, motivated in part by the Pittsburgh Downtown Plan (Strada, 2009), there was a noticeable change in the supply of available parking spaces. This is largely a result of the modified zoning ordinances that relaxed the minimum parking requirement for prospective properties. These measures were informed by the need to promote better land use within the Golden Triangle. Apart from these measures, a host of initiatives, both demand and supply side approaches have also been employed in addressing the parking situation. Demand initiatives include the use of education, technology-based solutions, leasing changes and a pricing strategy that is meant to incentivize workers and the growing number of downtown residents to patronize fringe garages. This increased utilization of peripheral parking facilities is expected to free up parking spaces in the core for individuals who come downtown for business and retail purposes and demand parking for a much shorter period. Supply side projects include the new PNC building that will provide 300 parking spaces for users; planned construction, pending government approval, of public space garages within the Cultural District, 5th Avenue and Grant Street Corridor; and the North Shore Connector that will facilitate access to parking spots on the North Shore.

However, significant attrition in parking supply is expected in light of present developments. The reduction in parking supply will come from new construction, the repurposing of existing parking facilities and the movement of corporate entities. These developments include the Penguin's hockey arena uptown, the movement of Equitable Resources to the Dominion Tower and the expansion of the University of Pittsburgh Medical Center (UPMC) downtown; a revitalized riverfront trail that is expected to reduce the supply of available parking spaces; the new Consol Energy Center that became operational in 2011; Point Park's academy village initiative that will bring architectural and streetscape improvements to the university's downtown neighborhood; and a reduction in the number of surface lots either for development purposes or for speculative reasons. These attritions, coupled with the expected five-fold increase (University of Pittsburgh, 2010) in the downtown's area resident population by the end of the decade creates a situation that may further increase the already high downtown population per parking space ratio. In addition, vehicular traffic is expected to grow in the downtown area given the increasing relevance of downtown magnets like the PCT and the Davis Lawrence Convention Center and the continuing cuts to public transit.

Stakeholders' Analysis and Needs Assessment

It is in light of the aforementioned challenges that a project on smart parking emerged. The provision of real-time and predictive information on parking was mooted to address the concerns of stakeholders who were dissatisfied with the current parking situation in downtown Pittsburgh. Information on the stakeholders including their expectations and objectives and program measures that address these objectives are provided in Table 1.

Table 1: Stakeholder Analysis and Measures

Stakeholder	Expectations & Objectives	Measures/Indicators
PCT/PCT Patrons	Decrease in late coming incidence to events	Percentage of patrons arriving after curtain
	Reduction in parking search time	Difference in pre/post average search time
	Reduction in search time variability	Mean parking search time deviation
	Reduced parking stress for patrons	Perception measures through surveys
Garage Owners & Management	Improved management of lease holders	Predicted demand estimates
	Increased flexibility in handling contingencies	Predicted demand estimates
Pittsburgh Downtown Partnership (PDP)	More positive perceptions on parking	Perception measures through surveys
	Reduced cycling time	Average parking search time outcome measure
	Decreased greenhouse gas emissions	Estimated value of congestion reduction

A needs assessment that builds on the stakeholders' expectations and objectives was conducted. Data collection was done through an in-person survey that was administered to patrons who were attending a Pittsburgh Cultural Trust event. In all, a total of 736 individuals were surveyed about their perceptions on parking within the Cultural District in the time period between September 18th, 2010 and January 23rd, 2011. *Parking search time* is measured by the mean search time and this measure is also used as a proxy for *search time variability* given that the search time data has an exponential distribution. Parking satisfaction captures the respondent's parking experience for the specific day the survey was administered while the overall parking experience averages her parking experience in the past 3 months. Baseline data obtained with respect to the key objectives are presented below.

Table 2: Baseline Data on Key Program Objectives

Program Objectives	Data
Parking search time	7.3min
Late coming incidence	27.0%
Perception about parking (% *indicates those surveyed without a positive response*)	
Parking satisfaction	25.7%
Ease of finding a parking space	22.4%
Overall parking experience	22.7%

For respondents who drove or car pooled, 82% parked in garages, 14% on surface lots and only 4% made use of on-street parking. One out of every four respondents said they have been late for a Cultural District event because they had difficulties finding a parking spot. Approximately three out of every ten reported not having a positive experience with overall parking when coming for a Cultural District event. The needs assessment also revealed that patrons are reluctant to use fringe parking lots partially because of security issue and because of the long walk during the winter season with seven out of every 10 persons surveyed revealing their preference for closer proximity as compared to lower price.

The smart parking application is intended to address these issues. The application's real time component provides visitors information on available parking spaces, thus, reducing the need to cycle and removing the need for a trial and error parking availability search process. The predictive model allows both visitors and garages operators to plan ahead and equips them to better manage parking demands. For example, garages can increase the supply of available parking spots by making provision for valet parking assuming a predicted higher demand for parking spaces and visitors may opt to use the public transit or a park and ride option in instances where predicted demand outstrips supply within the district. The subsequent systems development section documents the features of the application that make these goals a reality.

Systems Development

ParkPGH falls within the realm of initiatives involving information technology that are often called *smart parking* solutions. These initiatives are broadly categorized into two broad sections: *parking guidance systems, and real-time vs. prediction information*. The former is concerned with the design and behavioral responses to *parking guidance systems (PGS)* which use variable message signs (VMS) to inform drivers about available parking spaces. Much of the literature on parking guidance systems is concerned with transit and park-and-ride lots (Shaheen & Kemmerer, 2008) and examples of the actual implementation of this system exist as describes in Orski (Orski, 2003). Another stream of work on PGS explores their use inside of parking facilities ((Caicedo) and (Caicedo)). 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 only available through mobile devices, interactive voice response (IVR) and the Internet.

The latter - *real-time vs. prediction information* - examines 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, et al., 2007) and Teng et al. (Teng, et al., 2008) are examples of the former that provide parking prediction models based on information exchanged between wirelessly connected vehicles for use *during* a trip. The ParkPGH prediction model is an event-based parking prediction model for use *before* a trip begins. The prediction model uses 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, et al., 2004). For example, an individual coming downtown to watch a performance could establish, with some degree of certainty, the probability of finding a parking spot assuming the Pittsburgh Penguins are playing the same evening. Drivers may subsequently incorporate these parking predictions into their pre-trip planning.

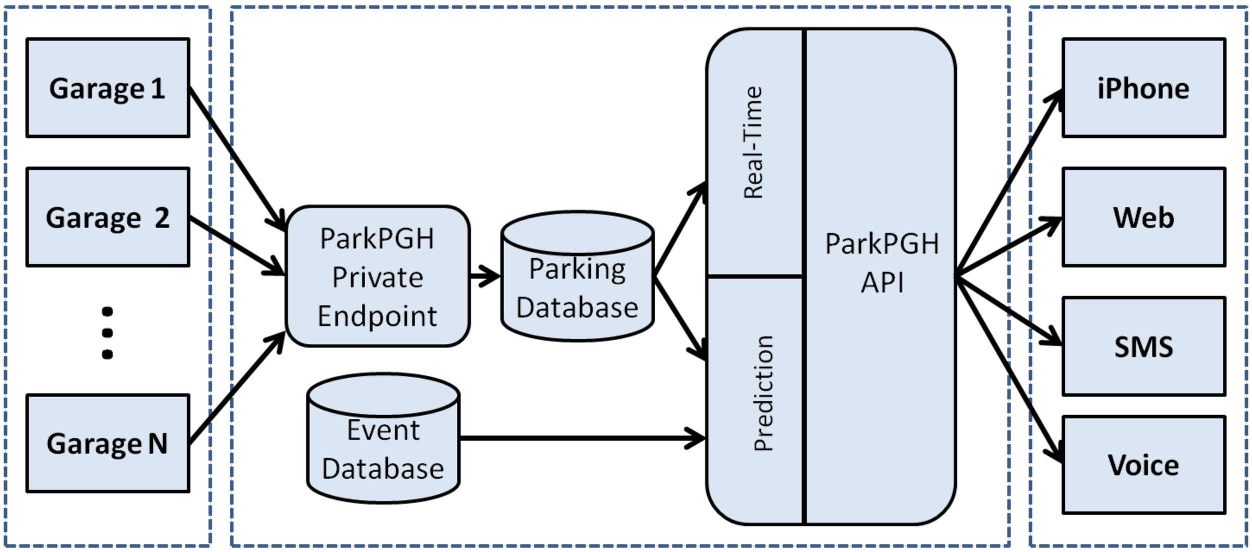


Figure 1: The ParkPGH System Diagram with interfaces for both garage operators and end users.

To implement this system, we employed an innovative approach that combines system development and integration with a parking prediction algorithm as shown in Figure 1. 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 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 input historical parking and event data occurring 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 prong technological innovation was deployed through a pilot program that monitors eight parking garages totaling 5000 parking spaces, representing approximately 20% of the total parking supply in downtown Pittsburgh and over 90% of the total parking supply in the cultural district. Parking information is updated every minute and is delivered through channels that include websites, iPhone app, SMS text, voice and a mobile version of the website that provides the same information as the traditional website but optimized for mobile devices such as Blackberries and Android phones.

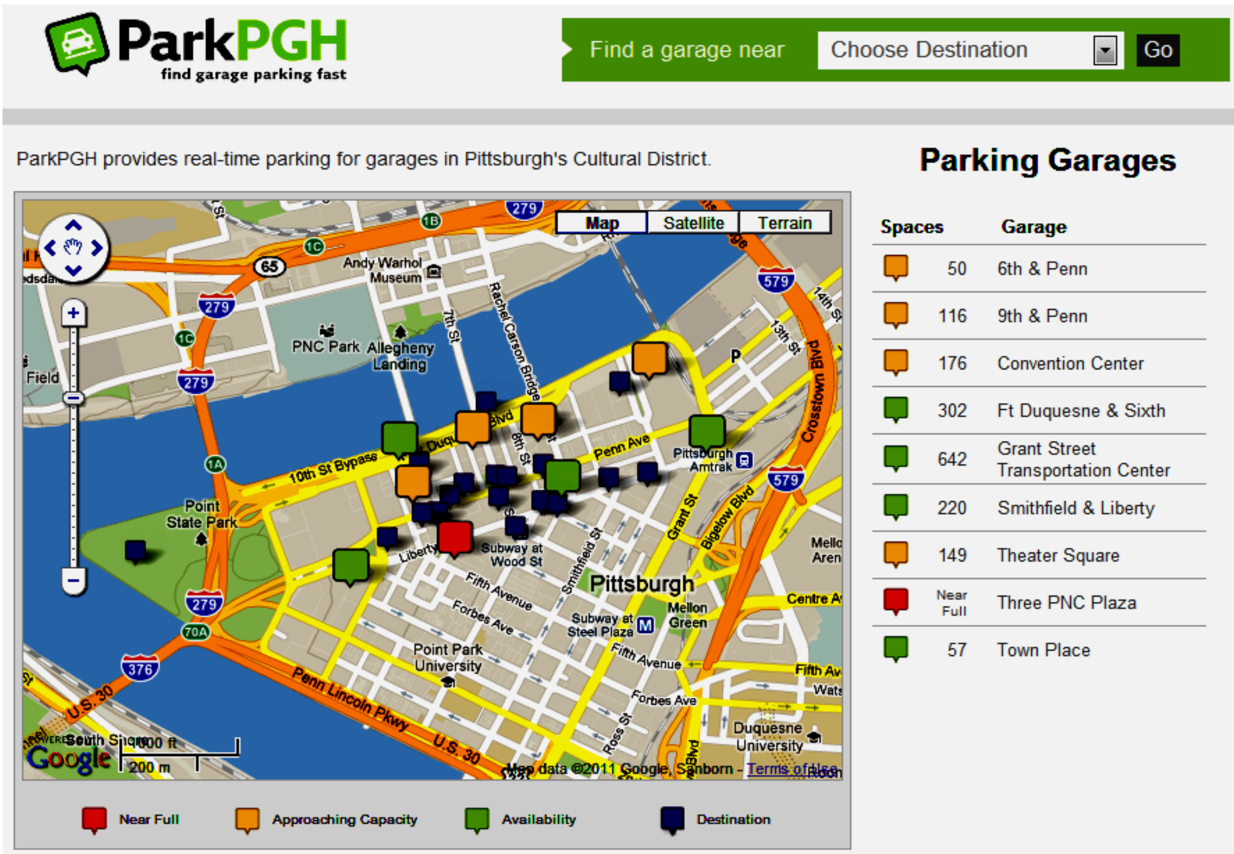


Figure 2: ParkPGH website displays Pittsburgh's downtown map with available parking spaces

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 2.

ParkPGH makes parking availability information available on a variety of distribution channels. The iPhone app and the mobile website feature a scrollable view, listing each available parking facility and its parking space availability. Clicking on a garage reveals more information, including the facility address, map and pricing. In addition to parking garage information, popular destinations are displayed so that visitors can locate their targeted destination and find the closest available parking. A mobile version of the website, *m.parkpgh.org*, provides the same information as the traditional website but will be optimized for mobile devices, focusing on the Blackberry and iPhone. The reader may observe that the exact available number of parking spots is not provided when the garage is deemed nearly full. Double clicking on any one of the garages produces more detailed on the garage chosen. Information provided includes rates and the theaters in close proximity to the garage available space. For SMS, visitors can text PARKING to 412-423-8980 to obtain a list of downtown parking lots in the order of availability. Similar to the SMS offering, visitors may call 412-423-8980 to receive a list of parking lots with available space. A text-to-speech system will announce the parking lot names and the percentage of available spaces for each lot. The call can be interrupted at any time by pressing the number of the parking lot (1, 2, 3. etc.) to get location and pricing information for the respective lot.

Figure 3 is a screenshot of the ParkPGH iPhone application. It shows both the real-time and prediction capabilities of ParkPGH. 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 2 events occurring near Theater Square garage that influences future parking availability. This predictive capability is a distinguishing feature of ParkPGH.

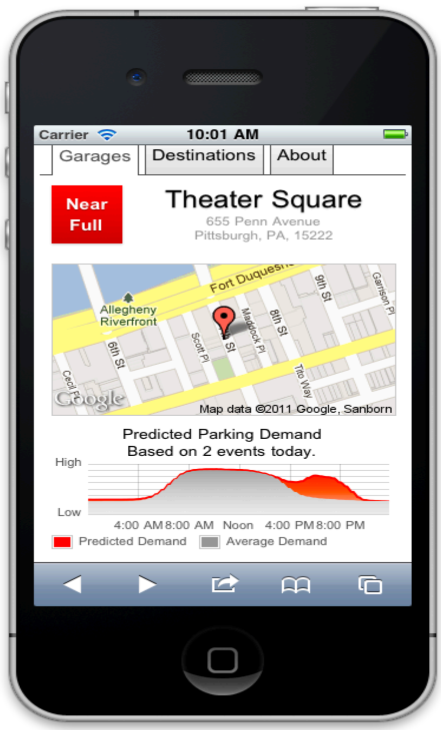


Figure 3: The ParkPGH iPhone displays the output of the Parking Prediction model.

Predicting Parking Availability

The underlying theoretical model for the number of parked cars is a non-stationary multi-server queue with no waiting space. The state of the system is the number of available parking spaces in the garage. We make no assumptions on the distribution of parking times. Our key assumption is that the time varying arrival rate is a function of the events occurring in the vicinity of the garage. We are able to observe the arrivals to the parking garage when it is not full. This is accomplished by collecting the electronic counts from the garage entrance gate technology. We also observe the events occurring in proximity of a given garage, in addition to the weather. We use these two data sources to predict the number of available parking spaces, and whether the garage is full or not full.

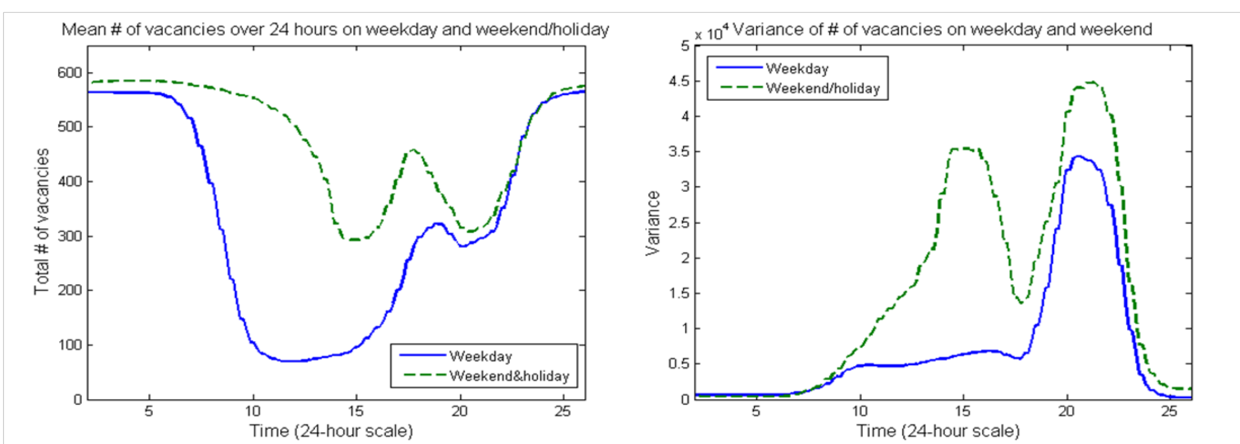
ParkPGH was developed for a long term prediction user scenario. That is, our prediction algorithm focuses on parking availability several hours or days in advance. Thus, the class of prediction models we consider does not use the current number of vehicles parked, or any information about the number of parked vehicles throughout the day. However, time of day is available and critical for our models. The design rationale behind this decision is that the real time information provided by ParkPGH is useful for short term prediction, while the prediction algorithm is useful for longer term scenarios. Currently, ParkPGH does not provide predictions for availability for the minute to hour time scale. Such a model is the focus of future enhancements to ParkPGH.

We conducted the prediction analysis using a range of methods. The prediction models are presented in two parts. First, we focus on predicting the number of available parking spaces at a given time from the set of events and weather data using neural network based predictors. Also, given the fact that the driver directly cares about whether the garage is full or not full, we have provided a robust approach that reduces the possibility of Type II errors. 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 classifier, classification and regression trees (CART) and a neural network, form a complement to the continuous prediction methods.

Data Description

We describe the model for Theater Square Garage which is one of the eight garages participating in ParkPGH. The analysis for the other garages is identical. The training parking data included the number of available parking spaces for every 10 minute interval for 24 hours (144 data per day) from 11/9/2008 to 7/10/2010 (609 days or 87 weeks). Figures 4a-b shows the average available parking spaces for weekdays and weekends/holidays and their corresponding variances. A huge drop in the number of available spaces observed between 10am and 3pm on weekday is considered to be work related because the number of spaces is more or less stable (low variance), where we can use a historical data to predict the available spaces. In contrast, a drop at around 3pm on weekend and drops at around 8pm on weekend and weekday are considered to be event related because the number of spaces fluctuates heavily (high variance) depending on the occurrences of events. In fact, variances have three clear peaks at 3pm on weekend and at 8pm on weekday and weekend. It is precisely these periods that accurate parking information is needed most.

The prediction model estimates parking vacancy based on events such as theater



Figures 4a-b: Mean and Variance of the number of vacancies on weekday, weekend and holidays

performances and sports games held in Pittsburgh's Cultural District and the weather conditions using rain and snowfall measured in inches The model uses data from the Pittsburgh Cultural Trust detailing events in the Benedum Center, Byham Theater, O'Reilly Theater, Heinz Hall, Pirates baseball game, Steelers football game, University of Pittsburgh football game, and Penguins ice hockey game. We split all events into 3 categories - 1) morning (before noon), 2) day (12:10pm-4pm), and 3) night (after 4:10pm). The events are subsequently used as predictors in estimating the number of available parking spots using neural networks. Other predictors include a vector set of regressors that capture the weather condition – snow and rainfall measured in inches and dummies for the day of the week and days that are declared as holidays. The descriptions of variables employed for the prediction model are provided in Table 3.

Table 3: Definition of Variables

Variable	Definition
Categorical Independent	
Event Dummies	
Theater - ben2, ben3, byh2, byh3, or2, or3, hnz2, hnz3	Dummies for theater events at the Benedum Center (ben*); Byham Theater (byh*); O'Reilly Theater (or*); and Heinz Hall (hnz*)
Sport - pir2, pir3, hnzf2, hnzf3, pen2, pen3, stl2, stl3	Dummies for sporting events - Pirates (pir*); Heinz Field (hnzf*); Penguins (pen*); and Steelers (stl*)
Day of the week - sunday, monday, tuesday, wednesday, thursday, friday, saturday, holiday	Dummies for day of the week
Time of the day - period	Specific period of the day measured in 10 minute increment
Numeric Independent	
Weather - snow, rain	Snow and rainfall measured in inches
Numeric Dependent	
avail	Number of available parking spots

The variables in Table 3 are classified broadly into three categories – categorical independent variables; numeric independent variables; and the numeric dependent variable. The morning, afternoon and evening split were used to capture the time when theater and sporting events were scheduled. Dummies for morning events were excluded since no event were scheduled within that time frame. In addition, given that both the Steelers and the University of Pittsburgh football program use the Heinz field, we used different dummies to make a distinction between these two events. Finally, we used a holiday dummy to isolate the different parking demand profile for any given holiday.

Neural Network

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) and the multi-layer feed forward net. Summary findings from the analysis using Palisade® Neural Tools is provided in Table 4.

Table 4: Summary of Neural Network Analysis

Summary	
Net Information	
Configurations Included in Search	GRNN, MLFN 2 to 6 nodes
Best Configuration	GRNN Numeric Predictor
Training	
Number of Cases	24418
Training Time	0:56:01
Number of Trials	31
Reason Stopped	Auto-Stopped
% Bad Predictions (30% Tolerance)	9.2309%
Root Mean Square Error	56.44
Mean Absolute Error	21.89
Std. Deviation of Abs. Error	52.02
Testing	
Number of Cases	6105
% Bad Predictions (30% Tolerance)	11.4005%
Root Mean Square Error	60.84
Mean Absolute Error	25.28
Std. Deviation of Abs. Error	55.33

The best net search out of a total of six different architectures utilized is the generalized regression neural network (GRNN) numeric predictor. The table shows the independent category variables, the independent numeric variables and the dependent variable which is the number of parking spaces available in the garage. In all 31 trials were carried out both for the training and the testing sample sets. Measures of how close the predicted values are to the eventual outcomes are provided in 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 have provided measures for the percentage of bad prediction. These indicators show the number of cases in the set for which the network predicted an output value that is statistically different from the actual known value. A bad prediction figure of 11.4% was obtained for the testing case, an indication that, on average, approximately 1 out of every 10 predictions will be wrong.

Figure 5 provides measures of the sensitivity of the GRNN predictions to changes in the nine most influential regressors. The cumulative value of the impact of all the regressors is normalized to 1 and the proportionate contribution of each of the regressor is provided as relative variable impact value. The higher the value for a given variable, the more it affects the predicted values for the dependent variable. The explanatory variable that represents the specific time of the day in 10minutes increment has the most impact followed by variables that capture the weather situation. This result is specific to a given net. The priority ordering of the regressors may be different for another net whose learning procedure allows it to discover higher significant contributions to predictions for a regressor that may have only a marginal impact in other nets.

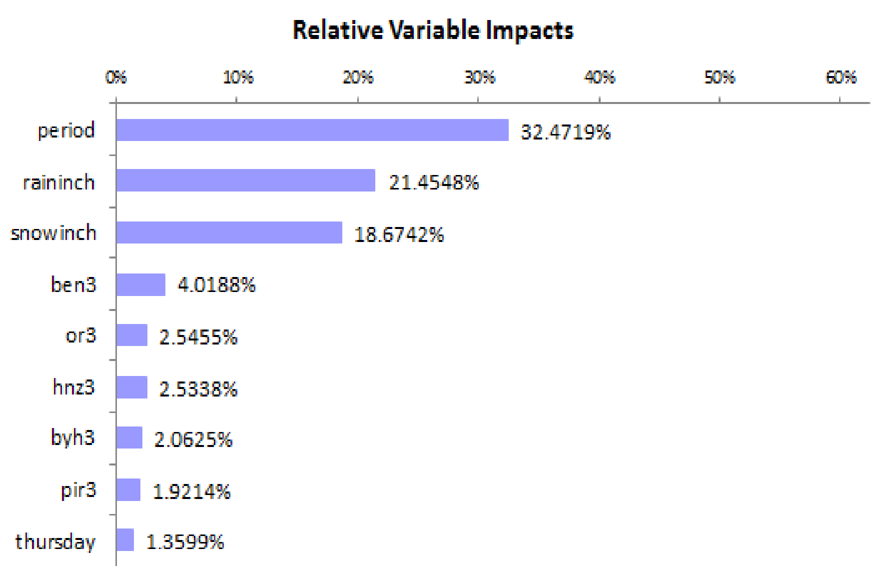


Figure 5: Relative Variable Impacts

Measures of Predictive Accuracy

Table 5 provides a comparative analysis of predicted accuracy measures using RMSE values. In all, 6 nets were trained and tested to identify the best one. The configuration of the multi-layer feed forward (MLF) net include the use of multiple nodes, ranging from 2 to 6 for the hidden layers. As could be seen, the GRNN performed best of all the nets trained and tested. The degree of improvement in the accuracy of predictions is especially noteworthy when the RMSE value for the best net is compared to that of the linear predictor.

Table 5: RMSE Values for different nets

Best Net Search	
	RMS Error
Linear Predictor	170.68
GRNN	60.84
MLFN 2 Nodes	144.62
MLFN 3 Nodes	136.96
MLFN 4 Nodes	139.81
MLFN 5 Nodes	136.85
MLFN 6 Nodes	138.72

Beyond aggregate measures of predictive accuracy, the summary table for the neural network analysis provides estimates of the percentage of bad predictions – specific instances where the actual outcome differs from the predicted value. We have enriched this measure by looking at the residuals and fitting the values into a distribution function. The ideal distribution is a Laplace distribution with mean value of -0.31 and absolute mean deviation value of 35.75. Finally, we have carried out sensitivity analysis to determine the reliability of our predictions –measured by the range of RMSE values and the ideal amount of data that should be set aside for the testing case. Results from the sensitivity testing have been 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 is 20% and the RMSE range from a low of 60.78 to a high of 62.90 for this threshold as shown in Figure 6.

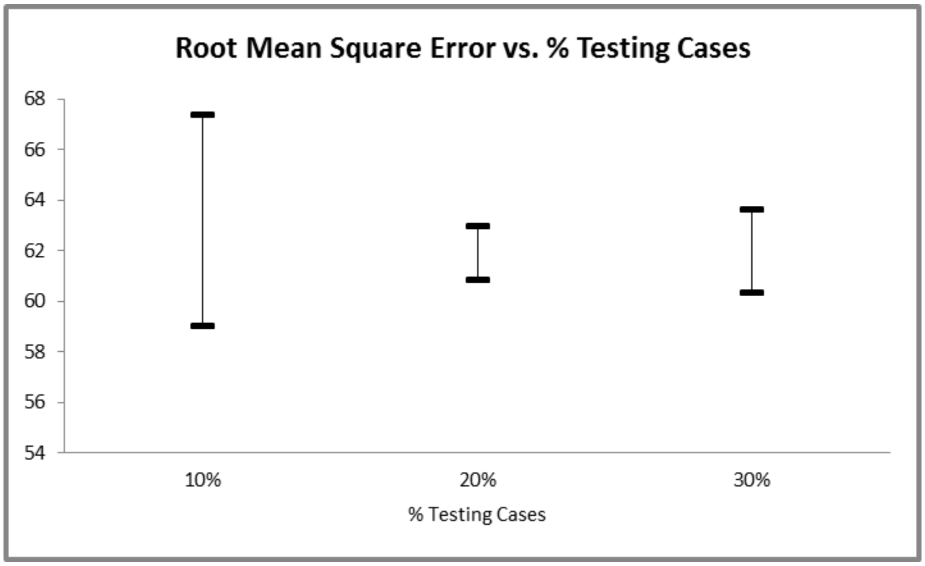


Figure 6: Mean Square Error bars and Measures of Dispersion across Multiple testing Cases

Rationale for the use of Predictive Classification Methods

Following up on the previous section, we intend to ascertain if the errors in the predictions could be localized to specific subsets of the possible realization of the dependent variable. Figure 7 addresses this. The 45^0 line represents the locus of points where the residuals are zero and the predicted and the actual values are of the same magnitude. Observations much further away from this line are indications of bad predictions.

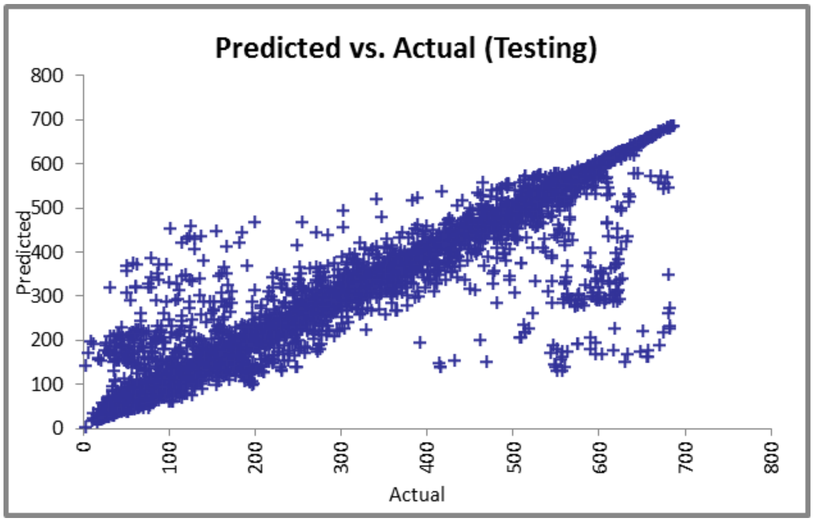


Figure 7: Scatter plot of Predicted vs. Actual (Testing)

From mere eyeballing, it is obvious that the net provides a better fit in situations where the utilization of the garage parking spots is neither low nor high. However, the predicted errors seem to propagate at the extremes. The predicted values were systematically overestimating the actual values for low values and consistently underestimating the actual values for high values. This explains the 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. The threshold level for dichotomizing the dependent variable was chosen as to be conservative and adjusted to ensure that Type II errors are avoided.

Predictive Classification Methods

Thus, as a compliment to the continuous variable prediction models, we explore classification methods based on machine learning. In order to accomplish the classification, we use a binary dependent variable, **full** or **not full**. The garage is considered full when the availability is less than 15% of capacity. This definition of **full** is consistent with the user interface developed for

ParkPGH. A user of the iPhone app or mobile website is shown that the garage is full if the current availability is less than 15%.

We use the same data set as for the availability prediction which contains 36,949 **not full** observations, representing 94% of the total observations, and 2409 **full** observations. The independent variables remain the same as in the continuous variable prediction models. We report on the prediction results of two fundamental machine learning classifiers: naïve bayes, classification and regression tree (CART). Additionally, we report results of a logistic regression, and a neural network

The naïve bayes classifier predicts the class **{full, not full}** using a an approach similar to an empirically driven maximum likelihood estimator. The key assumption is that the features, i.e. events, are conditionally independent given the class, **full or not full.** The classification and regression tree (CART) method is a nonparametric method which uses a binary tree on the features to classification and prediction. The number of branches and leafs are iteratively selected to minimize the square error of the prediction. The exact description of this iterative technique is beyond the scope of the paper, see (Loh, 2011) for more information.

When the dependent variable takes discrete values in a class, a logistic regression is natural. In our case the class is binary, **{full, not full}.** The functional form of the relationship between the dependent variable and independent variables is assumed to follow the logistic function. The model is typically estimated via maximum likelihood methods. The performance of these models are obtained using 10 fold cross validation. The results average root mean squared error, precision, and recall are reported in table 6. The root mean squared error (RMSE) measures the squared prediction error of each observation in the out of sample test set. The 10-k RMSE is the average root mean square error over 10 cross validation test sets. It is worth pointing out 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.

Table 6: Summary of Results for the Classification based Predication Methods

	10-K RMSE	Precision	Recall	Binary Outcome
Naïve Bayes	0.2519			
		0.951	0.964	**not Full**
		0.299	0.237	**Full**
CART	0.1293			
		0.984	0.991	**not Full**
		0.842	0.756	**Full**
Logistic	0.2219			
		0.942	0.995	**not Full**
		0.449	0.059	**Full**
Neural Network	0.15517			
		0.9818	0.9928	**not Full**
		0.6041	0.8487	**Full**

Precision is the ratio of the number of true positives divided by the sum of the number of true positive and false positives. *Recall* is the ratio of the number of true positives divided the number of true positives and false negatives. A false negative for class **not full** is when the garage is **not full,** but is classified as **full.** A false negative for class **full** is when the garage is **full,** but is classified as **not full.** As seen in Table 6, the performance of CART is superior to the other measures. This is not surprising given that naïve bayes and the logistic regression assume the features are independent. CART on the other hand does not make that assumption and is free to build tree to exploit any correlation in the feature structure. The resulting CART has 341 nodes with 171 leaf nodes. Each of the leaf nodes correspond to a unique combination of features or scenarios.

The CART extension to the GRNN numeric predictor has provided a robust prediction platform. The hybrid approach mimics the real time feed that is currently running across multiple channels of the smart parking application where 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 will be 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 as regards the tolerable level of error.

Management Solutions for Multiple Stakeholders

One of the key challenges we encountered in implementing ParkPGH were the problems created by the unique environment within which the smart parking application is deployed. The parking facilities featured in the pilot program are owned and operated by entities with different management structures. The fragmented ownership and diverse management structure make it extremely difficult to design a standard approach that will be amenable to all the garages. When the project was conceptualized, it was thought that there was a uniform method of determining the number of currently available parking spots in the garages, along with a way of determining when the garage could be identified as being "full." However, each parking garage has its own "culture" of determining how and when to identify the garage as being "full." Variables that factor into that decision include the number of leased spots to hold open, use of valet parking, the threshold level at which the "full" sign goes up and garages that distinguish between hard and soft full.

This lack of standardization has made for significantly increased complexity in the algorithms used in the ParkPGH application. We have addressed this, in part, by developing a novel web portal for garage managers. This platform allows the documentation of lease management strategies and process issues that shape the idiosyncratic features exhibited by some garages. The information is shared with the software development team with the objective of exploring the possibility of these garage specific traits to be taken into consideration when the smart parking application is being fine-tuned. Secondly, we have accommodated the subjectivities emanating from different management structures through the level of granularity of information provided. An example is the decision to suppress the information on the number of parking spaces available when the garage is deemed full or close to maximum capacity. The possible options by which information is relayed to the public were pilot tested to ascertain the ideal level of detail especially when garages are close to full capacity.

Application's Impact

In order to improve upon and ascertain the value added by ParkPGH, a series of indicators were tracked. 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, number of unique views, number of automated phone responses, bounce rate and average duration of page views. A sample of the measures tracked is shown below.

Figure 8 shows the number of daily requests for selected delivery channels between January 1st 2011 and July 31st, 2011. The usage volume is typically higher during the weekdays compared to

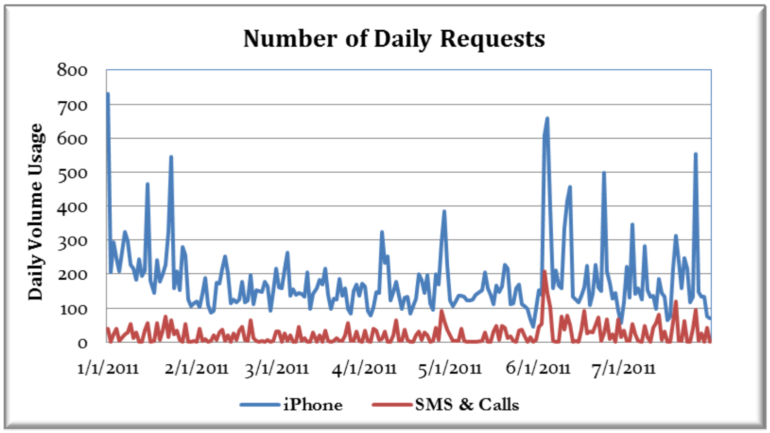


Figure 8: Daily requests for iPhone and SMS shows consistent usage of ParkPGH.

weekends except when events are scheduled. For example, the noticeable spike in usage on the weekend of June 3rd to the 5th is attributed to the Pittsburgh JazzLive International, a weekend of music that includes outdoor stages, visual art shows, musicians of international repute and a JazzLive crawl.

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.

Outcome measures that document the impact of ParkPGH are provided in Table 7. 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 individuals reporting a 4-6 minute reduction in search time being in the majority.

Table 7: Outcome related measures

DOCUMENTED IMPACT	%
ParkPGH has made finding parking spaces easier	
% of respondents with positive response	57.2
Specific reduction in search time	
% of respondents reporting a reduction in search time	48.6
% of respondents with 1-3min reduction	17.1
% of respondents with 4-6min reduction	22.9
% of respondents with more than 6 min reduction in search	8.6

Conclusion

ParkPGH is a smart parking application that uses parking and event data to provide real-time and predictive information on the availability of parking in eight parking garages within the Pittsburgh Cultural District. The system has proven effective in reducing search time when finding a parking space. The reduction in search time has led to less cycling, changed patrons' perception as regards the parking situation downtown and has made the Golden Triangle a more attractive destination both for business and pleasure. In 2011, more than 300,000 unique inquiries were made of the ParkPGH application.

The initiative employs an innovative approach that combines systems development and integration with a parking prediction algorithm. The prediction utilizes both the GRNN and the CART method in making parking predictions. This approach has enabled the development team achieve a high degree of prediction accuracy. In addition, the evaluation of the program implementation provided information that allowed modifications to be made to the smart parking application.

The conceptualization and the execution of the program design was a carefully planned process that includes inputs from key stakeholders. Engaging stakeholders in the initial phase and in program implementation has provided a platform robust enough to handle deviations from accepted norm. Some of these challenges were created by the unique environment within which the smart parking application was deployed. In contrast to many cities where all the city parking facilities are owned by a single entity, the parking facilities featured in the ParkPGH pilot program are owned and operated by a wide variety of entities with vastly different management structures. These issues were addressed in part, by developing a novel web portal for garage managers. This platform allows the documentation of lease management strategies and process issues that shape the idiosyncratic features exhibited by some of the garages.

Acknowledgements

We would like to thank Marc Fleming and John Mumper of the Pittsburgh Cultural Trust for guidance and Merrill Stabile, Jim Funovits and Don Levkus from Alco Parking for providing us with the parking data set, events calendars and many valuable suggestions. The Pittsburgh Parking Authority and the Pittsburgh Downtown Partnership were also instrumental in the success of ParkPGH. We also benefitted from the comments and suggestions of the editor and two anonymous referees. Finally, we would like to acknowledge Deeplocal Inc. for building the ParkPGH application. This work was supported by the Benter Foundation and a Carnegie Mellon University initiative called Traffic 21.

References

Bos, I., Ettema, D. & Molin, E., 2004. Modeling Effect of Travel Time Uncertainty and Traffic Information on Use of Park-and-Ride Facilities. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 1898, pp. 37-44.

Caicedo, F., 2009. The use of space availability information in 'PARC' systems to reduce search times in parking facilities. *Transportation Research Part C: Emerging Technologies,* 17(1), pp. 56-68.

Caicedo, F., 2010. Real-time parking information management to reduce search time, vehicle displacement and emissions. *Transportation Research Part D: Transport and Environmen,* 15(4), pp. 228-234.

Caliskan, M., Barthels, A., Scheuermann, B. & Mauve, M., 2007. *Predicting Parking Lot Occupancy in Vehicular Ad Hoc Networks.* s.l., Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, pp. 277-281.

Loh, W., 2011. Classification and regression trees. In: *Data Mining and Knowledge Discovery 1.1.* s.l.:Wiley Interdisciplinary Reviews, pp. 14-23.

Orski, K., 2003. Best Space Scenario. *Traffic Technology International,* pp. 54-56.

Partnership, P. D., 2012. *State of Downtown Pittsburgh 2012.* [Online] Available at: http://www.downtownpittsburgh.com/SOPD [Accessed 23 May 2012].

Shaheen, S. & Kemmerer, C., 2008. Smart Parking Linked to Transit: Lessons Learned from Field Test in San Francisco Bay Area of California. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 2063, pp. 73-80.

Strada, L., 2009. *Redeveloping downtown pittsburgh – the last 20 years.* [Online] Available at: http://www.stradallc.com/files/07.28.09_white_paper_1.pdf [Accessed 24 May 2012].

Teng, H., Qi, Y. & Martinelli, D. R., 2008. Parking difficulty and parking information system technologies and costs. *Journal of Advanced Transportation,* Volume 42, p. 151–178.

Teng, H., Qi, Y. & Yi, P., 2002. Parking Search Time and Information Identification for Off-Street Parking in New York City. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 1800, pp. 44-52.

University of Pittsburgh, P., 2010. *Pittsburgh: A New Portrait.* [Online] Available at: http://upress.pitt.edu/htmlSourceFiles/pdfs/9780822943716exr.pdf [Accessed 24 May 2012].