Stay Elsewhere? The Economic Impact of Location-based Hotel Features: A View from Remote Sensing Image Analysis

Anindya Ghose **Panagiotis Ipeirotis** **Beibei Li**
Stern School of Business, New York University

1. Research Question

One of the common Web searches that have a strong local component is the search for hotels. Customers try to identify hotels that satisfy particular criteria, such as food quality, service, and so on. A crucial feature is the location of the hotel. For example, everything else being equal, a hotel close to the beach is typically more desirable than a hotel that is separated by a highway from the beachfront. Similarly, a hotel located in the downtown area can charge higher prices than that located at the outskirts of a city, and still make a sale. Such location-based features like "proximity to the beach" or "near downtown" represent important characteristics that influence the desirability of a particular hotel, and in turn can influence the corresponding rate of that hotel. However, currently there are no established economic metrics that can isolate the economic impact associated with these various features of a local hotel. Our goal is to empirically estimate the economic value of different hotel features, especially the location-based features given the associated local infrastructure. We aim to do so by combining state-of-the art econometric modeling with spatial data and image classification methods. Then, after inferring the economic significance of each feature, we will incorporate these features in a ranking function to improve the local search for hotels. This will result in a real-world application of our research.

2. Approach

In general, we aim to identify how particular hotel features will affect the desirability and price of a particular hotel. Based on these results, we improve the quality of local search for such hotels. We have formulated a novel dataset based on the characteristics of 9463 different hotels which are currently located in the United States.

More specifically, our research work involves three stages. First, we aim to identify the particular set of hotel features that are most highly valued by customers and hence, contribute to the aggregate prices of the hotels. We detect these features based on an anonymous survey from 100 independent common customers by using the "Amazon Mechanical Turk (MTurk)" (http://www.mturk.com), which is an online tool providing a web service API to integrate computer artificial intelligence directly into application by making requests of humans. By analyzing the replies to our proposed question, we have selected the most frequently mentioned features for hotels. Depending on the different point of view the hotel features represent, these hotel features are then divided into three categories: (a) location-based hotel features, for example, "proximity to the beach" or "near car rentals"; (b) price-based hotel features, for example, "discount" or "rewards"; (c) service-based hotel features, for example, "wireless Internet".

In the second stage, we have collected all the corresponding features for each hotel. We have been collecting these data in different ways over the past three months. The price-based and service-based hotel features have been crawled from TripAdvisor.com. For the location-based features, which comprise the larger component of our work, we have conducted our study based on the analysis of online remote sensing image data. These data have been retrieved by utilizing Microsoft Virtual Earth Interactive SDK (available at http://dev.live.com/virtualearth/sdk/) through the Visual Earth Tile System. The data have been extracted as 256*256 hybrid satellite images containing both road and aerial information, corresponding to all the 9463 different US hotel venues with 4 different zoom levels for each. However, all the location-based feature information cannot be easily derived from the images in the same way. Such feature like "proximity to the beach" is easily recognized by machine learning methods, whereas results may not be so satisfying for "near car rentals". Our approach is hence conducted at a comprehensive level, involving a combination of machine learning, computer artificial intelligence as well as human intelligence techniques. Specifically, we have been collecting the location-based hotel features primarily through three different methods: (1) Features providing rich texture information in the images, such as "proximity to the beach" and "near downtown" have been derived by Image Classification. (2) To infer features that contribute to less texture information but to more commercial connections, such as "located near car rentals", we have used the Virtual Earth Interactive SDK. (3) For features that are hard to recognize by any of the artificial intelligence algorithms but are easier retrieved by humans, such as "located near a highway", we have used the MTurk online survey.

In the final stage, we compute the economic value associated with each hotel features and measure the extent to which they impact the pricing decision for a local hotel business. We perform hedonic regressions[1, 2] to decomposes a hotel into all three groups of its constituent features, and measure the contribution of each hotel feature to the overall hotel rate. In this way, we are able to estimate the economic value of those qualitative features, especially location-based features like "near downtown", which have not been examined by prior researchers.

3. Main Findings/Expected Contributions

Our research allows us to not only quantify the economic impact of hotel features, but also by reversing the logic of this analysis, it allows us to identify the most crucial location-based features that influence the desirability of a particular venue. By incorporating the value of these features in the local result ranking function, we can provide customers with the best-value hotels early on, hence improving the quality of local search for such hotels. Meanwhile, for the hotel business owners, our analyses will also strengthen their ability to impute the optimal pricing strategies.

4. Current Status of Manuscript

We have finished the data preparation, including hotel feature identification in stage one and feature collection in stage two. We have run some preliminary analysis with the experimental dataset. We are now ready to start our final evaluation. By next March, we expect to have our completed paper presented at the conference, if given the opportunity.

References

[1] Rosen, S. Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy* 82, 1 (Jan.-Feb. 1974), 34-55.

[2] Feenstra, R.C. Exact Hedonic Price Indexes. *Review of Economics and Statistics (*1995), 77:634-654.