Effect of Retail Layout on Traffic Density and Travel Distance

Nicholas Ballester, Bradley Guthrie, Stephanie Martens, Corinne Mowrey, Pratik J. Parikh, Xinhui Zhang

Department of Biomedical, Industrial and Human Factors Engineering
Wright State University
Dayton, OH 45435 USA

Abstract

Past research in the field of retail facility layout has been limited to evaluating customer travel behavior assuming a fixed store layout or evaluating store layout assuming fixed traffic density zones. This study relaxes these assumptions and explores the effect of department layout on traffic density and shopper travel distances in a retail store. The analysis consists of three stages. First, actual shopping data is used to identify the probability distributions of (i) number of items in a trip and (ii) departments where these items are located. Second, for a predefined department layout and a given list of items, all shortest paths that a shopper could opt for in the store are determined. Finally, overall aisle densities upon aggregating these travel paths are calculated. Our preliminary results suggest that departments with a high percentage of items purchased do not necessarily need to be located at the rear of the store in order to draw a reasonable amount of traffic there. Moreover, layouts that achieved a close-to-uniform traffic distribution also had lower average shopper distances. Weighting the interior of the store produces the most uniform traffic distribution.

Keywords
Retail Store Layout, Traffic Density, Department Location, Travel Distance

1. Introduction

Warehousing facilities are an intermediate location in the supply chain, whereas retail facilities are the end of the supply chain. They directly interface between the product and the customer. Thus, a retail store has the potential for substantial impact on revenues, for better or for worse. In this, the customers (hereafter referred to as shoppers) are the most vital part, yet the store may not have direct control over them. The focus, then, must be on influencing the shopper as much as possible towards improving sales, without drastically affecting their shopping experience.

Much research has been performed with this goal in mind, ranging from marketing and promotions to product placement and shelf allocation to the psychology of typical shopper behavior. One important factor to be considered is the layout of the store and its effect on shopper travel through the store, which in turn affects product exposure. This holds especially true in the case of impulse purchases, which are a key component of store profits. Impulse purchases can be influenced to some extent by the store, whereas pre-determined purchases from a shopping list are much more difficult to affect.

Retail facility layout may well be the most important factor of retail profits. Shopper traffic flow within a store is a more effective indicator and a stronger influencer of store sales than volume of shoppers in the store [1]. Consequently, improving store layout would increase profits much more than trying to attract more shoppers through marketing and promotions, which consume considerable amounts of time, money, and other resources.

Clearly, the retail facility layout problem is quite different from the warehousing facility layout problem. Unlike warehousing, retail layout is usually driven by profits, not by travel distance [2]. Unfortunately, while the warehouse facility layout problem is a well-defined and studied problem in IE/OR, the retail facility problem is not as delineated, despite the large impacts than can be made in this area due to the economic size of the retail sector. This deficiency can be largely attributed to the fact that shopper behavior in a retail store is difficult to analyze and control due to the high level of variability, whereas in the warehousing sector picker travel paths can be defined and controlled.
Current research in retail store layout has two main limitations. Researchers have either evaluated shopper behavior and store traffic flows assuming a fixed store layout (department locations and aisle configurations), or they have evaluated different store layouts (i.e., department locations) assuming fixed traffic patterns and density zones. In reality, shopper traffic patterns through a store are most likely influenced to some extent by the store layout and department locations. In this study, the effect of department location on traffic density and travel distance is examined, by varying store layouts with dynamic shopper travel behavior in a simulation-type process. The resulting traffic patterns are then analyzed in an attempt to summarize the interactions between department location and density and distance.

2. Literature Review

Many studies have examined shopper behavior from a more psychological perspective — how do shoppers think, feel, and act (or react) under different store circumstances. One example is an analysis performed by Harrell et al. [3] for the specific case of crowding in the store. It was an empirical study that examined shoppers’ perceptions, actions, and satisfaction. Another study looked at several key metrics, including perceived usefulness, ease of use, entertainment, and time spent, for various standard store layouts such as the “grid” and “racetrack” [4]. The context was online shopping—web designs for retail sites that were constructed using adaptations of the principles of the physical store layouts in question.

A large number of descriptive studies of shopper travel paths drawn from empirical evidence (usually RFID tracking of shopper carts) have been performed. Larson et al. [5] used such a study to analyze shopper behavior and dispelled the myth that shoppers travel systematically up and down aisles. They determined that shoppers usually make quick excursions into and out of an aisle, using the perimeter of the store as a “racetrack” or main thoroughfare (this leads to the importance of end-of-aisle displays). Hui et al. [6] applied the same approach in a different vein and examined typical shopper behavior in comparison to optimal (Traveling Salesman) paths. Yada [7] attempted to expand this technique by using character string analysis in order to allow for the consideration of such aspects of shopper behavior as the amount of time a shopper spent in a section and typical section-to-section transitions. More recently, Takai and Yada [8] used this method to examine the effect of the time of day on shopper behavior.

The earliest analytical approach to the retail facility layout problem was performed by Farley and Ring [9], who in 1965 developed a complex stochastic model for predicting shopper behavior in terms of the probability of a shopper transitioning from one area or department of a store to another. This model was tested on several layouts and found to be fairly accurate. However, the authors did not discuss traffic densities within the store, or the effect the store layout has on these traffic densities.

In 2004, Peters et al. [10] introduced a set of preliminary approaches to the process of modeling the facility layout problem, such as evaluating a layout, characterizing store traffic, and assigning departments. They also examined expected tour lengths for given shopping lists for several standard store layouts: aisle-based, hub and spoke, and serpentine. Building on this research, in the following year Botsali and Peters [11] developed a model for optimizing a serpentine layout using a network approach. In their model, they dealt directly with product visibility and impulse purchases.

A different approach was taken by Yapicioglu and Smith [2] in developing a mathematical model for the optimal placement of departments within a retail store in order to maximize store revenue and satisfy desired adjacencies. Although their model was quite complex, it was limited to a racetrack store layout and it assumed fixed traffic density zones. All analysis was at the departmental level; the shopper did not enter the equation directly.

This study distinguishes itself from the previous research in that it deals directly with store layout in an analytical fashion and does not examine shopper behavior from a psychological perspective; it is not merely descriptive, but theoretical and predictive; it is not limited to fixed layouts (e.g., serpentine or racetrack), fixed shopper travel paths, or fixed traffic density zones; it analyzes the effect of varying department location on traffic density. It is, however, limited in that it does not directly address product visibility and impulse purchases, and maintains a high-level view of the layout problem (aisle structure, shelf allocation, and other more detailed aspects of the retail facility problem are not considered). It is an exploratory study into an area of the retail facility layout problem which has not yet been addressed.
3. Methodology

3.1 Data
A local store for a Fortune 500 retailer was selected for this study (for proprietary reasons, the name of the retailer is not disclosed). Figure 1 below demonstrates the physical store layout, which was used as the initial model and reference point in the rest of the analysis. As can be seen, this particular store follows a racetrack-style layout.

To simplify the analysis, some of the broad product categories in the store were grouped together and treated as a single “department.” This was done following the schema already in use by the retail store for their own analysis purposes. The resulting thirteen departments (see Table 1, at right) were then used throughout the rest of the analysis.

Table 1 (at right) also shows the percentage of items (in terms of quantity, not revenue) that were purchased in each of the thirteen departments in the store. Snacks was the largest category, accounting for almost 17% of all items purchased in the store, while alcohol was the smallest at less than 1%. These percentages were calculated from 815 separate transactions totaling 551,441 items purchased.

The number of items per shopper was computed using transactional data from the retail store. As seen in Figure 2 below, the distribution of the number of items per shopper was found to follow a Poisson distribution with mean (λ) equal to 8.429. The sample size for this data was 22,278 unique shoppers.
3.2 Network Representation
In dealing with store layouts, a 3x3 network representation with nodes and edges was used (see Figure 3, below), in which nodes represented an area of the store and edges represent travel pathways between these areas. The physical locations of the departments in a store were mapped to the nodes in this representation in order to identify paths to be followed by the shopper and calculate corresponding distances. This methodology allowed for the approximation of a complex store layout into a manageable network for analysis, and permitted any layout to be analyzed simply by changing the department-to-node mapping. In order to approximate travel through the central area (e.g., pharmacy to meat/seafood), all travel passing through the central area was aggregated into two orthogonal (interior) cross-aisles of the network (i.e., edges BEH and DEF in Figure 3). The final result of this department-to-node mapping for the physical store in question can be seen in Figure 4 below.

3.3 Procedure
The method followed in this study was composed of three main phases: generating a shopping list, calculating shortest path, and estimating traffic densities. We discuss each of these phases in detail below.

3.3.1 Generate Shopping List
A number of items was first assigned to each shopper according to the Poisson distribution determined from the data from the retailer in question. Each item was then assigned a product category corresponding to one of the 13 departments (as defined in Section 3.1 and not necessarily according to the product categories of the physical floor plan) within the store. The likelihood of a product category/department being chosen for an item depended on the

![Figure 2: Number of items purchased per shopper](image)

![Figure 3: Network representation of a retail store](image)

![Figure 4: Department-to-node mapping for the selected store](image)
percentage of items purchased from that department, which was calculated from the data. A shopper was allowed to have multiple items from the same department. Given these departments, a final list was assigned in terms of the nodes the shopper was required to visit in order to fulfill their purchases, irrespective of the number of items at each node. The necessary nodes were determined by the department-to-node mapping scheme, which in turn was determined by the layout selected.

For example, consider Shopper 1 who was assigned 7 items, one each in departments 1, 3, 10, and 11, and three in department 9. Based on the department-to-node mapping scheme (see Figure 3 and 4) for the approximation of the real-world retail store, Shopper 1 would need to visit nodes A, B, D, F, and G in order to fulfill his/her list.

### 3.3.2 Calculate All Alternate Shortest Paths

Once a list of nodes to visit had been assigned to each shopper, the next phase of the analysis was to calculate all possible shortest paths for the shopper. The list of nodes to be visited by each shopper was broken down into a series of two-node pairs, and for each two-node pair all equally optimal paths through the network were computed. All combinations of each of these two-node paths were then considered, and those with the shortest overall distance were retained. In calculating distances, the edges were assumed to be equally weighed, and each was assigned a distance of 1 unit for this analysis. The actual distance can be obtained by simply multiplying the resulting distance with an appropriate distance measure corresponding to a store.

To continue with the example introduced above, in order to visit nodes A, B, D, F, G, and then exit the store, Shopper 1 could travel by the path A-B-C-F-E-D-G-D-A or A-D-G-H-E-F-E-B-A. Both these paths resulted in the shortest distance of 8 units. In this example, there were 12 possible alternate (shortest) paths that Shopper 1 could choose from to visit the nodes on his/her shopping list. We identify each one of them and consider them as equally likely (i.e., a weight of 1/12 in this case) for our subsequent calculations.

### 3.3.3 Estimate Traffic Density

The final phase of the analysis was to calculate the traffic density between the nodes and construct a density flow map for the layout. Traffic density along an edge of the network was quantified by \( \omega \), the percentage of the total traffic contained along the edge between two nodes. All equally optimal paths for each shopper were considered when calculating the \( \omega \) for each edge, but they were weighted according to the number of paths that shopper had available.

### 3.3.4 Layouts Analyzed

These three phases were repeated for eleven different store layouts, represented by mapping the departments to different nodes. One thousand shoppers were generated and assigned shopping lists, and then all of them were run through each layout.

Table 2 gives the descriptions of the different layouts. A policy similar to turnover-based in warehousing was used to construct layouts 2, 3, 7, and 8 by locating departments such that the store was divided into zones with differing percentages of items purchased, concentrated either towards the entrance of the store (front-loaded), the rear (rear-loaded), the corners (corner-weighted), or the interior (interior-weighted), respectively. In layouts 4, 5, and 6, a department arrangement was sought that would achieve a uniform traffic density throughout the store (in an attempt to distribute the store-wide traffic and ensure each area of the store was equally exposed to the shopper). In layout 4, departments were allocated in order to equalize the percentages of items purchased at each node as much as possible. Layouts 5 and 6 were modifications (attempted improvements) of other layouts which had achieved fairly uniform traffic in most, but not all, edges. Departments were randomly allocated to nodes in layouts 9, 10, and 11, which were analyzed for comparison purposes.
3.4 Assumptions

Our results are specific to the store layout and shopping data we obtained. However, the methodology developed in this study could easily be adapted by changing the department-to-node mapping, changing the location of the entry/exit node or adding additional entry nodes, changing the mean or the type of the distribution of number of items per shopper, and changing the percentages of items purchased in each department. We also assumed that the transactional data we provided, and which we used to determine the distributions for number of items per shopper and items per department, contained ‘planned’ purchases by the shopper and that these purchases remain unaffected by layout changes. In reality, shoppers may make several ‘unplanned or impulse’ purchases that may be reflected in the data we were given; however, it was difficult to distinguish them from planned purchases. These unplanned or impulse purchases are likely to be most affected by layout changes, and methods to estimate them need to be designed.

In calculating shopper paths, we assumed that shoppers traveled along the shortest paths, did not backtrack or sidetrack, and required only a single visit to a department to purchase all items from that department. In reality, shoppers will often deviate from optimal paths due to forgetting to pick one or more items from that department (resulting in backtracking), impulse purchases, and curiosity. However, it has been demonstrated that, overall, shoppers often travel by near-optimal paths — “…shoppers tend to pick up their purchased products in an order close to that suggested by the TSP but tend to depart from the shortest point-to-point path (i.e., travel deviations) as they move through the store” [6].

The physical area occupied by the departments was not considered in the department-to-node assignments, but this was mitigated by the low ratio of departments to nodes (13 to 8) and by the constraint that no more than three departments could be assigned to a single node.

4. Results and Discussion

The following figures (5-10) display the results of several of the more noteworthy layouts. The edges are shown with their corresponding omega values. The percentage of items purchased at each node is also displayed.

4.1 Traffic Density Distribution

Among the 11 layouts analyzed, the traffic densities along the edges directly adjacent to the entrance/exit node (A) were stable and always higher than the density along any of the other edges; values ranged between 11-14%. This makes intuitive sense as the aisles near the entrance of a store would have high traffic. But it was surprising to find that the percentage of items purchased allocated to the nodes at the ends of these edges did not affect the density excessively, even if it was disproportionately higher or lower than in another layout. As a general rule, edges towards the interior and the rear of the store were more susceptible to changes in store layout, while edges towards the front of the store were more robust. The edges at the farthest corner of the store (across from the entrance) experienced the most fluctuation in traffic density, ranging from 4% to over 11% in the most extreme cases (the front-weighted (Figure 8) and rear-weighted (Figure 7) layouts, respectively).

A tendency towards traffic patterns that approximated a racetrack was noticed. In general, edges along the perimeter of the store had a higher percentage of traffic than edges in the interior of the store. Moreover, across the majority of the layouts tested, the edges at the rear of the store maintained a reasonable percentage of traffic relative to the rest of the edges, indicating that the common practice of locating highly-frequented departments at the rear of the store may be unnecessary. In fact, edges in the interior of the store usually suffered from lower traffic in this study, more so than the rear, so the interior of the store may be the area on which to focus such efforts.

4.2 Uniform Traffic Density

Uniform traffic density was very difficult to achieve, and no layout was found in this study that resulted in all edges having an equal percentage of store traffic. Given the fact that the edges adjacent to the entrance always monopolized shopper traffic (between the two of them, they accounted for 25% of store traffic on average), a perfectly uniform distribution of traffic density could not be expected. However, it was possible to approach a uniform distribution among the rest of the edges. Weighting the interior nodes (the five nodes other than the corner nodes) provided the best results—it most closely approached a uniform distribution of traffic density throughout the store (see Figure 9, above). In general, any layout in which percentages of items purchased were concentrated towards the interior nodes performed better than others in terms of uniformity of traffic distribution. Another example of this was the real-world layout, which was weighted more towards the interior than the periphery, and
which was fairly uniform in traffic density (Figure 5). In contrast, allocating departments so as to evenly distribute the percentages of items purchased among the nodes as far as possible did not achieve uniform density—rather, it resulted in a racetrack traffic pattern (see Figure 6, above). Weighting the corners achieved an even more pronounced racetrack pattern (Figure 10).

Figure 5: Approximation of real-world layout (Layout 1)

Figure 6: Near-equally distributed items (Layout 4)

Figure 7: Rear-loaded (Layout 3)

Figure 8: Front-loaded (Layout 2)

Figure 9: Interior-weighted (Layout 8)

Figure 10: Corner-weighted (Layout 7)
4.3 Potential Tradeoff between Traffic Density and Travel Distance
In order to explore the correlation between traffic density and travel distance, we attempted to identify measures that would quantify the uniformity of a layout (in terms of traffic density). It is interesting to note that while mean (first moment) provides little value in determining uniformity, variance (second moment) is useful as it captures the spread of traffic density values across the edges. However, variance does not indicate if the densities are symmetric. We, instead, used skewness (third moment) to estimate at an aggregate level the degree of symmetricity (or lack thereof) in traffic densities on the edges. Here we define skewness as the concentration of density across 4 zones in the store layout, beginning with the entrance (Node A) and ending diagonally opposite at the back of the store (Node I). Zone 1 consists of edges AB and AD; Zone 2 consists of edges BC, BE, DE, and DG; Zone 3 consists of edges CF, FE, EH, and HG; and Zone 4 consists of edges FI and IH. For a perfectly uniform layout, cumulative densities would result in \((Z_1, Z_2, Z_3, Z_4) = (16.67, 33.33, 33.33, 16.67)\), perfectly symmetrical about the midpoint. Skewness values were calculated using the expression,

\[
S = \sqrt{n} \frac{\sum_{i=1}^{n}(X_i - \overline{X}_{avg})^3}{\sum_{i=1}^{n}(X_i - \overline{X}_{avg})^2}^{\frac{1}{3}},
\]

where \(S\) is skewness, \(X_i\) is the value of each observation, \(\overline{X}_{avg}\) is the mean of the observations, and \(n\) is total number of observations. The skewness of a perfectly uniform layout would equal 0.

Figure 11 shows that as the skewness and shopper travel distances are negatively correlated (correlation coefficient = -0.67). That is, as the values for skewness of various layouts increase, the corresponding average shopper travel distances decrease. For instance, the front-loaded layout (L2) has the highest skewness value and the lowest travel distance, while the back-loaded layout (L3) has a relatively low skewness value and high travel distance. Intuitively this makes sense; as farther parts of the store are being reached, shopper travel distance increases.

![Figure 11: Plot of Skewness vs. Average Travel Distance](image-url)

Although the skewness measure is able to capture the lack of symmetry between the traffic density among various zones (aggregated over edges), it has its limitations. In particular, the measure is unable to clearly depict how traffic is moving to the back of the store (i.e., through exterior or interior edges). For example, layouts 3 (rear-loaded), 4 (equally distributed items), and 7 (corner-weighted) exhibited low skewness values, but these layouts resulted in a markedly racetrack traffic pattern (Figures 6, 7, and 10). Interestingly, layout 8 (interior-weighted) and even layout 1 (real-world) also exhibited relatively low skewness values, but these layouts approached nearly-uniform traffic conditions.
density (Figures 5 and 9). However, the average shopper travel distances acted as a differential; i.e., layouts with nearly-uniform traffic patterns tend to exhibit lower travel distances than layouts with racetrack traffic patterns.

5. Summary and Future Research
The retail facility layout problem is a sparsely researched area of IE/OR. What research does exist is limited by one of two assumptions: either traffic density is fixed or the store layout is fixed. By relaxing these two rules, this study was able to analyze the effect of varying department locations on store traffic density and shopper travel distance. Using a network approach to represent a store layout and estimate distances in a simulated study using real data, several interesting, albeit preliminary, results were obtained. Among these, the more notable observations are:

- Departments with a high percentage of items purchased do not necessarily need to be located at the rear of the store in order to draw a reasonable amount of traffic there.
- Locating departments such that the percentage of items purchased is evenly distributed throughout the store does not necessarily result in an even distribution of store traffic; rather, it leads to a racetrack pattern.
- While a perfectly even distribution of store traffic is not possible due to the monopoly of traffic held by the entrance area, it is possible to approach a close-to-uniform traffic density throughout the rest of the store. In this respect, layouts with percentages of items purchased weighted towards the interior perform the best.
- Layouts that achieved a close-to-uniform traffic distribution also exhibited lower average shopper distances. Distances were the greatest in layouts resulting in racetrack traffic patterns.
- Skewness of traffic density across zones is negatively correlated with average travel distance, confirming that placing popular departments at the rear of the store increases travel distance.

The implication of some of these observations can be substantial. From a store manager’s perspective, maximizing shopper travel distances will lead shoppers deeper into the store ensuring that they spend more time in the store and are able to view more areas/products, both of which could increase the likelihood of impulse purchases. In contrast, uniform traffic density ensures that, on average, shoppers visit nearly every section of the store giving the store manager an increased area to display promotional items. From a shopper’s perspective, a shorter average distance would seem desirable, due to less walking required to purchase items on their shopping list. A front-loaded layout would be better in this case due to lower average travel distance, but this may lead to overly crowded areas, which may not resonate well with a good shopping experience that many shoppers tend to seek. Loss of sale and, eventually, loyalty may ensue. In contrast, a layout with uniform traffic density will spread the shoppers throughout the store resulting in less crowding and an improved flow. Leading the shoppers deeper in the store may cause dissatisfaction and crowding in those areas; have we not experienced crowding in the dairy section of a store which is typically in the rear part of the store? Trading off traffic density and travel distance seems critical to both the store manager’s and shopper’s perspectives.

As future research, it would be interesting to analyze stores with different distributions of items per shopper and items per department, and stores with additional entrances or an entrance/exit point located at a different part of the store. Even for the case of the selected layout, the experiment could be rerun with a more accurate department-to-node mapping, by introducing more nodes, and splitting the thirteen departments into smaller categories. Additionally, more variations of this real-world layout can be examined to verify if the tendencies noticed are indeed trends and to examine more closely some of the more intriguing observations.

References