ABSTRACT
It is well-established that increased product visibility to shoppers leads to higher sales for retailers. In this study, we propose an optimization methodology which assigns product categories and subcategories to store locations and sublocations to maximize the overall visibility of products to shoppers. The methodology is hierarchically developed to meet strategic and tactical layout planning needs of brick-and-mortar retailers. Layouts in both levels of planning are optimized considering eligibility requirements and complete set of shopper paths, thus, they successfully capture the unique shopping behaviour of consumers in a store’s region. The resulting mathematical optimization problem is recognized as a special instance of the well-known Quadratic Assignment Problem, which is considered computationally as one of the hardest optimization problems. We adopt a linearization technique and demonstrate via a real-world numerical example that our linearized optimization models substantially improve the store layout, hence, can be used in practical applications as a vital decision support model for store layout planning.

1. Introduction
The retail industry is globally an essential sector which is known for slim profit margins and fierce competition. Grocery retail is additionally dynamic as it must answer to ever-changing customer shopping behaviours and business ecosystem. Therefore, grocers increasingly employ retail analytics to make strategic decisions in well-known processes such as pricing, assortment, marketing, and promotion planning. In addition to these traditional areas, store layout optimization can provide a retailer significant edge against the competition. The present work focuses on store layout optimization to maximize visibility of products to shoppers as they navigate the store pick products. In this article, we loosely associate product “visibility” and “exposure” and use the terms interchangeably.

The problem foundationally stems from the industry adage “unseen is unsold.” Although grocery shoppers mostly rely on predetermined shopping lists, unplanned purchases (also known as impulse purchases) can constitute 30–50% of their purchases (Kollat & Willett, 1967; Mishra & Mishra, 2010). Unplanned spending is triggered by in-store stimuli such as visual confrontations with a product (Rook, 1987). Visual stimuli can be enhanced mainly by (1) increasing the shelf space allocated to products or (2) intelligentlylocating products in the store so that customers are exposed to maximum number of products along their shopping paths. Dreze et al. (1994) established that increased product visibility leads to higher impulse purchases. The first approach of visual stimuli was employed in Xu et al. (2021) where floor space was optimally allocated to product categories such that the total store revenue is maximized. Related to the second approach, Hui et al. (2013) assessed the hypothesis that traveling farther in a store results in higher unplanned spending. They statistically showed in a field experiment that as shoppers travel farther from their planned path, they make up to 50% more unplanned purchases on average. In grocery industry, practitioners have long employed some tactics to lengthen shoppers’ travel in the store. A famous example of such tactic is “hide milk at the back of the store.” These expertise-based strategies support layout decisions, yet, the final layout might be far from optimal because they do not determine the optimal layout based on how customers pick the products in their baskets. In this work, we optimize the store layout via maximizing the total exposure of product categories and subcategories which are calculated over all customers’ transactions.

The contribution of this article is multi-fold. We first identify the shopper path-based maximum product exposure problems within the framework of the typical store layout planning processes for different planning horizons and formulate mathematical models for them. We further develop a
mathematical programming-based solution approach to solve the computationally intractable instances of the problems. We finally validate the algorithm using real customer transaction data and store layout location data and derive insightful outcomes. The rest of this article is organized as follows. Section 2 summarizes relevant literature on retail layout optimization. Section 3 outlines the problem and the data. In Section 4, we develop mathematical optimization models for the store layout with maximum exposure problem required to support strategic level and tactical level planning as well as an integrated strategic-tactical planning model. Next, we present linearized versions of the proposed optimization models in Section 5. Section 6 reports on our computational experience using a real-world planning instance from a European grocery chain. Finally, we summarize our contribution and outline some future directions of research in Section 7.

2. Relevant literature

In current practice, there are more design considerations to store layout than optimization to boost impulse buys. A store layout should provide plausible shopping experience through store atmosphere and in-store traffic patterns to enhance store loyalty and sales. Several researchers have shown the positive relationship between store layout and consumer purchase intentions (Ainsworth & Foster, 2017; Baker et al., 1992; Merrilees & Miller, 2001). Also, store layout should be aligned with in-store operations such as shelf stocking to provide operational efficiency (Lewison, 1994; Vrechopoulos et al., 2004). Consequently, well-designed store layouts appeal to the customers and contribute to both product sales and store profitability (Cil, 2012). In the present work, we aim to satisfy these desired layout characteristics and rules by introducing them as constraints in our mathematical formulation.

Vrechopoulos et al. (2004) identified three major layouts in conventional retail settings: grid, free-form, and racetrack. In grocery retail, we commonly encounter stores combining the grid and racetrack aisles as shown in Figure 1. The parallel arrangement of long aisles in the centre of the store forms the “grid” which is circumvented by a “racetrack” with shelves leaning to the sides and back of the store. In this formation, main pathway on the perimeter branches into grid aisles, thus, it facilitates planned shopping behaviour with flexibility and speed in finding products from a shopping list (Larson et al., 2005; Levy et al., 2004; Lewison, 1994).

Retail facility layout optimization has only been in researchers’ radar in the recent decades. In the early work of Botsali and Peters (2005), different configurations of serpentine layout were analysed to determine one which maximizes impulse purchase revenues. Botsali (2007) furthered the analysis by including the grid layout. Cil (2012) built a customer-oriented supermarket layout using the product clusters derived from transactions. The layout is created to facilitate customer navigation by placing associated product clusters close to each other. The author claims that customer-appealing layouts would translate into higher purchase conversion rate. Conversely, Peng (2011) disperses the defined “must-have” products on a grid layout in an aim to maximize the sales coming from “impulse-buy” products as customers travel between “must-have” items. Yapicioglu and Smith (2012b) optimized a department store layout to maximize revenue with respect to the layout requirements. In the study, impulse purchases were reflected in revenues by scaling them with respect to the traffic densities of

Figure 1. Grocery store with location level layout for category assignment.
the store zones. The problem was solved with a Tabu Search algorithm on simulated problems of up to 20 departments on a racetrack structure. In their succeeding work, Yapicioglu and Smith (2012a) translated the previous problem into a bi-objective model where they developed layout requirements as another objective of the problem. Ozgormus and Smith (2020) also studied the similar bi-objective problem for a supermarket with racetrack aisles surrounding a grid. They simulated revenues on a function with predefined revenue coefficients and impulse purchase rates based on marketing and store manager’s input. The problem is solved for 25 groups of product categories with a Tabu Search algorithm. In another implementation for a grocery store layout, Flamand et al. (2016) maximized the impulse revenues by determining the optimal locations of 31 product categories on different layout configurations. In their related work, Flamand et al. (2018) incorporated product assortment affinity rules in the layout problem and optimally allocated 100 product categories in a grocery store with 30 shelves.

There has been a focus in the recent literature on product exposure due to rack orientation in the layout of retail stores. Mowrey et al. (2019) exploited the shopper’s visual experience with store racks and analytically estimated the exposure considering the human field of view and the orientation, height, and width of the racks. They showed that non-orthogonal orientation of racks can increase exposure by up to 250%. They also determined the optimal rack features that maximize exposure of rack facings in a store using particle swarm optimization (PSO) heuristics (Mowrey et al., 2018). Guthrie and Parikh (2021) further contributed to the estimation of rack exposure by introducing 3D field of view and extending the problem to include the rack curvature. In their subsequent work with an experimental study, they maximized impulse profit and determined the best rack layout orientation and curvature after accounting for floor space cost (Guthrie & Parikh, 2020). In addition to rack layout decisions, Karki, Guthrie, and Parikh (2021) jointly optimized rack features and product locations in the layout using a PSO algorithm, which resulted in improvements in impulse profits of about 10% in their experiments.

The aforementioned literature considers preset traffic densities as one of the drivers of the impulse purchases. However, traffic densities are calculated with respect to the locations of product categories in the current layout; these densities could potentially change with the optimized layout. Indeed, Ballester et al. (2014) showed that traffic densities ultimately change with the changes in product category locations. For instance, as a popular milk product is relocated in a supermarket, the associated traffic will follow it to the new location. To account for variable store traffic intrinsic to the problem, layout optimization can be carried out by considering the shopper paths. As mentioned before, several researchers showed that product exposure, and eventually sales, increase with the increase in the length of shopping paths (Hui et al., 2013; Inman et al., 2009; Khodol et al., 2010). Supported by this finding, Boros et al. (2016) aimed to increase the path length for the average customer in a supermarket. They optimized the store layout for 23 departments and 27 representative customer baskets. Longer travel paths presumably increase the product exposure, however, product exposure can vary for different sections of the store. For instance, in Figure 1, product exposure is higher at the back and centre of the store as products are displayed on both sides of the aisles, on the other hand, it is lower at the front of the store as products are only displayed on one side. Therefore, a better objective is to directly optimize the product exposure. Recently, Hirpara and Parikh (2021) maximized the impulse purchase revenue coming from product exposure on a customer’s shopping path. They solved the problem using a simulated annealing algorithm for a store with 20 departments using one representative basket.

3. Problem and data description

The current problem, i.e., to find an optimal store layout to maximize exposure of product subcategories, is a real-world application motivated by the store layout planning operations of a major grocery chain in Europe. It is studied on a typical medium-sized store located in an urban centre. The representative layout, a combination of racetrack and grid structures, is provided in Figure 1. On this layout, 20 categories on three main shelf groups are defined with respect to fixture and product category requirements. These are peripheral shelves (1–4), endcaps (5–12) and aisle shelves (13–20). Based on this typical layout, store layout planners make decisions first at strategic and then at tactical levels. Strategic level store layout planning happens every five to ten years during which stores are reset for a new store concept and refreshed look. Stores are remodelled around established customer shopping behaviours and most recent industry trends.

In our work, the strategic level layout model is termed as the Level-1 model and it translates into a problem of finding 20 optimal category locations. In this model, categories are only allowed to move within the same prescribed locations due to fixture
requirements. Optimal category locations from the Level-1 model form a basis for the tactical level layout planning problem which we address as a so-called \textit{Level-2 model}.

The Level-2 model is used to assign subcategories on a more granular layout as in \textbf{Figure 2} where a total of 48 subcategory sublocations are nested in the 20 category locations. At this level, a subcategory can only be assigned to a sublocation under its own category’s location from the Level-1 model. The Level-2 problem determines the 48 optimal subcategory sublocations within the already determined optimal category locations from the Level-1 model. The optimization problem based on the Level-2 model is solved on a quarterly to yearly frequency so that store layout adapts to seasonal product assortment changes while maintaining the familiarity of the store’s layout to the customers and keeping the operational costs of changing subcategory sublocation reasonable.

\textbf{Table 1} summarizes the assignment of product subcategories and their categories in the current store layout. In both the Level-1 and Level-2 problems, we aim to improve overall product exposures from the current store layout by optimally reassigning categories and subcategories to new locations to alter customers’ shopping paths.

The problems mentioned above are solved using data from 20,842 customer transactions. We assume customers pick categories in descending order of “popularity” which is defined as purchase frequencies derived from each customer’s transaction history. This “popularity” based assumption is analogous to the common approach where the layout problem is solved for a set of representative baskets. Additionally, in the Level-2 problem, a customer is assumed to pick all the subcategories belonging to a category in descending popularity order before proceeding to the next category. Customers are considered to travel the shortest path between locations in the Level-1 problem (sublocations in the Level-2 problem). While finding the shortest path, we use the centre-to-centre distance of the category locations and the subcategory sublocations in Level-1 and Level-2 problems, respectively. The walkway networks in \textbf{Figures 1} and \textbf{2} represent how customers travel within store for the corresponding problems. In each network, nodes are placed at the centre of every category and subcategory location, store entrance, and store exit. Then, distances between each location are pre-calculated using these nodes. Next, we calculate the total exposure for each shortest path. For every successive node pair on the walkway network, we pre-specify the exposure value, which is the number of sublocations exposed between the two nodes. We assume that a customer is exposed to a subcategory when passing by its sublocation. The total exposure of a path is the sum of exposure values for the successive node pairs constituting the path. In the situation where there are multiple shortest paths between two sublocations, we average the exposure values across the paths. Finally, a customer may get exposed to a sublocation multiple times while picking the order. In that case, we assume total exposure accumulates identically.

\section*{4. Model formulation}

In this section, we first present an integrated model that meets both the Level-1 (strategic) and Level-2 (tactical) planning requirements simultaneously. Due to practical reasons as well as the computational intractability of the integrated model, we then reformulate the integrated model into two sequential models: Level-1 and Level-2.

First, we define some notations for index sets, constants, and decision variables.
Table 1. Product categories and their locations in a grocery store.

| Location – Category | Fixture/Shelf Type | Subcategory (Sublocation) |
|---------------------|-------------------|---------------------------|
| 1 – Fruits          | Peripheral        | Bananas (1), Citrus (2), Hard Fruit (3), Soft Fruit (4) |
| 2 – Produce         | Peripheral        | Root Vegetables (5), Leafy Green Vegetables (6), Ready Cut Vegetables (7), Fresh Juice (8) |
| 3 – Frozen Food     | Peripheral        | Frozen Vegetables (9), Frozen Snacks (10), Frozen Pizzas (11), Frozen Potatoes (12) |
| 4 – Bakery          | Peripheral        | Bread Substitutes and Spreads (13), Fresh Bread (14), Prepackaged Bread (15), Signature Breads (16) |
| 5 – Fresh Dairy Food| Endcap            | Fresh Dairy Food (17) |
| 6 – Seasonal Deli   | Endcap            | Seasonal Deli (18) |
| 7 – Seasonal Non-food| Endcap         | Seasonal Non-food (19) |
| 8 – Seasonal Candy  | Endcap            | Seasonal Candy (20) |
| 9 – Fresh Dairy Drink| Endcap          | Fresh Dairy Drink (21) |
| 10 – Fresh Nuts     | Endcap            | Fresh Nuts (22) |
| 11 – Household Assortment| Endcap | Household Assortment (23) |
| 12 – Seasonal Confectionery| Endcap | Seasonal Confectionery (24) |
| 13 – Appetizers     | Aisle             | Mediterranean Food (25), International (26), Seafood Appetizers (27) |
| 14 – Condiments     | Aisle             | Salad Essentials (28), Sauces (29), Pickled Food (30) |
| 15 – Preserved Food | Aisle             | Preserved Vegetables (31), Preserved Meats (32), Sausages (33) |
| 16 – Chips          | Aisle             | Regular Chips (34), Premium Chips (35), Pop Corn (36) |
| 17 – Beer           | Aisle             | Bulk Beer (37), Bottle and Can Beer (38), Special Beer (39) |
| 18 – Water and Energy Drinks| Aisle | Energy Drinks (40), Flavored Water (41), Water (42) |
| 19 – Soda and Soft Drinks| Aisle | Coke (43), Soda (44), Ice Tea (45) |
| 20 – Soups and Eggs | Aisle             | Cakes and Cookies (46), Soups (47), Eggs and Non-perishable Dairy (48) |

**Index Sets:**

- **K**: the set of all locations with the first and the last locations representing the entrance and the exit of the store, respectively.
- **K**: the set of all sublocations with the first and the last sublocations representing the entrance and the exit of the store, respectively.
- **I**: set of product categories with the first and the last categories representing customer check-in (a dummy “product” representing a customer going through the entrance location) and customer check-out (a dummy “product” representing traversal through the exit location), respectively.
- **I****: set of product subcategories with the first and the last subcategories representing customer check-in and check-out, respectively.
- **G**: set of product subcategories that belong to category **i**.
- **L**: set of sublocations that belong to location **k**.

**Constants:**

- **E**: the pre-calculated exposure value (number of sublocations passed by on the shortest path) from sublocation **k** to sublocation **k**.
- **P**: the number of occurrences (over all customers) for which a customer walks to pick a product from subcategory **i** immediately after picking a product from subcategory **i**. Dummy “products” such as customer check-in and customer check-out are treated like any other products in this context.
- **A**: a binary constant with 1 indicating that product category **i** is eligible to be assigned location **k**. For example, customer check-in is the only product eligible to be assigned the first location (the entrance of the store), and the frozen food category must be assigned only to locations where freezers can be placed.

**Decision Variables:**

- **x**: the binary variable with value 1 if and only if product category **i** is assigned to store location **k**, and 0 otherwise.
- **z**: the binary variable with value 1 if and only if product subcategory **i** is assigned to sublocation **k**, and 0 otherwise.

With the above notations, we can now formulate the Integrated Model that covers both category assignments at the strategic level and subcategory assignments at the tactical level simultaneously.

**Integrated Model (IM):**

Maximize \[ \sum_{k_1, k_2 \in K} \sum_{i_1, i_2 \in I} E_{i_1, i_2} P_{i_1, i_2} x_{i_1, i_2} z_{i_1, k_1} z_{i_2, k_2} \] (1) subject to \[ \sum_{i \in I} x_{i k} = 1, \quad \forall k \in K, \] (2) \[ \sum_{k \in K} x_{i k} = 1, \quad \forall i \in I, \] (3) \[ x_{i k} \leq A_{i k}, \quad \forall i \in I, k \in K, \] (4) \[ \sum_{k_1 \in L_k} z_{i_1, k_1} = x_{i k}, \quad \forall i \in I, k \in K, i_1 \in G_i, \] (5) \[ \sum_{i_1 \in G_i} z_{i_1, k_1} = x_{i k}, \quad \forall i \in I, k \in K, i_1 \in L_k, \] (6) \[ x_{i k} \in \{0, 1\}, \quad \forall i \in I, k \in K, \] (7) \[ z_{i_1, k_1} \in \{0, 1\}, \quad \forall i_1 \in I^*, k_1 \in K^*. \] (8)
customers’ shopping paths. Note that both \( E \) and \( P \) values are constant and can be pre-calculated without relying on subcategory-sublocation assignment. For two given pairs of \((k_1, k_2)\) and \((i_1, i_2)\), the term \( E_{k_1k_2}P_{i_1i_2}z_{i_1k_1}z_{i_2k_2} \) represents the resulting exposure for the case that subcategory \( i_1 \) is assigned to sublocation \( k_1 \), and subcategory \( i_2 \) is assigned to sublocation \( k_2 \). Constraints (2) and (3) stipulate the one-to-one assignment requirement of a category to a location, more specifically, only one category can be assigned to a specific location, and only one location can be assigned to a specific category. Constraint (4) specifies that the assignment of a category to a location must be feasible subject to practical eligibility conditions. Constraint (5) and (6) establish further requirements in a Level-2 (tactical level) layout that a product subcategory must be assigned to a sublocation within the location that its corresponding category is assigned to. Finally, constraints (7) and (8) guarantee the binary nature of the decision variables.

The IM can be considered as a special instance of the Quadratic Assignment Problem (QAP) (Koopmans & Beckmann, 1957). The objective of the standard QAP is to assign a set of facilities to a set of locations in such a way that the total assignment cost is minimized. In our problem, we can view each product subcategory as a facility and construct a maximization instance of the QAP with additional restrictions imposed by the eligibility constraints (4), (5) and (6). QAP is a well-known NP-hard combinatorial optimization problem (Sahni & Gonzalez, 1976). Although extensive research has been done for over the last six decades, the QAP is still recognized computationally as one of the “hardest” optimization problems. No known exact algorithm can solve practical problems of sizes more than 30 locations in reasonable computational time. For example, Brixius (2000) reported an interesting computational experience in solving a 30-location benchmark QAP (Nugent 30 from QAPLIB) using a branch-and-bound method on grid computers. It required 1000 workstations over seven days to solve the QAP instance. Our preliminary computational experiments also confirmed that the integrated model with 48 sublocations/subcategories cannot be solved directly using a state-of-the-art quadratic programming solver such as Gurobi.

From a practical implementation perspective, as mentioned earlier, the strategic level store layout problem (with the Integrated Model as an instance) is solved infrequently (every five to ten years) and is only required for a store reset. On the other hand, the tactical level problem is required to be solved as often as quarterly to adjust the subcategory-sublocation assignments based on stationery category location assignments to accommodate seasonal adjustments to product assortment. This means that, from a practical standpoint, it is much more important for instances of a tactical level problem to be solved efficiently than a strategic level problem such as the Integrated Model. For this reason, we reformulate the Integrated Model into two distinct sequential models:

- **Level-1 Model**: Optimally assign categories to locations; and
- **Level-2 Model**: Optimally assign subcategories to sublocations based on given category locations from a Level-1 Model.

We define some additional constants associated with the Level-1 Model:

- \( E_{k_1k_2} \): the exposure (number of sublocations passed by on the shortest path) from location \( k_1 \) to location \( k_2 \). At this level, we assume all the sublocations of category is exposed when category is picked.
- \( P_{i_1i_2} \): the number of occurrences (over all customers) for which a customer walks to pick a product from category \( i_2 \) immediately after picking a product from category \( i_1 \).

Then the Level-1 Model can be formulated as follows:

**Level-1 Model (L1M):**

Maximize \[ \sum_{k_1, k_2 \in K} \sum_{i_1, i_2 \in I} E_{k_1k_2}P_{i_1i_2}x_{i_1k_1}x_{i_2k_2} \]

subject to \[ \sum_{i_2 \in I} x_{i_k} = 1, \quad \forall k \in K, \]

\[ \sum_{k \in K} x_{i_k} = 1, \quad \forall i \in I, \]

\[ x_{i_k} \leq A_{i_k}, \quad \forall i \in I, \quad k \in K, \]

\[ x_{i_k} \in \{0, 1\}, \quad \forall i \in I, \quad k \in K. \]

Letting \( \{x_{i_k}\}_{i \in I, k \in K} \) be an optimal solution obtained from L1M, we formulate the corresponding Level-2 Model as follows:

**Level-2 Model (L2M):**

Maximize \[ \sum_{k_1, k_2 \in K^*} \sum_{i_1, i_2 \in I^*} E_{k_1k_2}P_{i_1i_2}z_{i_1k_1}z_{i_2k_2} \]

subject to \[ \sum_{k \in K} z_{i_k} = x_{i_k}, \quad \forall i \in I, k \in K, i_1 \in G_i, \]

\[ \sum_{i_1 \in G_i} z_{i_k} = x_{i_k}, \quad \forall i \in I, k \in K, k_1 \in L_k, \]

\[ z_{i_k} \in \{0, 1\}, \quad \forall i \in I^*, \quad i_1 \in K^*. \]

We expect L2M to have higher objective function value than L1M. This is due to the fact that the
5. Linearization of the QAP

As described above, the store layout with maximum exposure problem can be formulated as an instance of the QAP and modelled as a standard mixed integer quadratic program (MIQP). It is well-known that the quadratic terms in the objective function can be eliminated and transformed into linear terms by introducing new variables and new linear constraints (Watters, 1967). After the linearization, the MIQP becomes a mixed (0–1) integer linear programming problem (MILP), so the existing methods and standard MILP solvers can be applied to solve the problem. Although, in general, MILP is considered as an “easier” problem than its counterpart MIQP, but the very large number of new variables and constraints increase the original problem size significantly leading to potentially considerable requirement in additional computing resources.

Researchers developed a variety of linearization techniques to transform an MIQP into a MILP depending on various methods to replace the quadratic terms. Some of the better known linearization methods include those proposed by Lawler (1963), Glover and Woolsey (1974), Kaufman and Broeckx (1978), Frieze and Yadegar (1983), and Adams and Johnson (1994). After experimenting with these linearization techniques, we adopt the Adams-Johnson linearization method from (Adams & Johnson, 1994) for the much reduced solution times of the resulting MILP formulation as compared to the original MIQP model.

For illustrative purposes, we show how we apply the Adams-Johnson relaxation method to L1M. The same procedure can be applied to IM and L2M directly to obtain associated MILPs. For each quadratic product term \( x_{i,k_1}x_{i,k_2} \) in the objective function of L1M, we introduce a continuous variable \( w_{i,k_1,k_2} \) with the following additional constraints:

\[
\begin{align*}
\sum_{i \in I} w_{i,k_1,k_2} &= x_{i,k_2}, & \forall i_2 \in I, & k_1, k_2 \in K, \\
\sum_{k_1 \in K} w_{i,k_1,i_2} &= x_{i,k_2}, & \forall i, i_2 \in I, & k_2 \in K, \\
w_{i,k_1,i_2} &= w_{i,k_2,i_1}, & \forall i_1, i_2 \in I, & k_1, k_2 \in K, \\
w_{i,k_1,i_2} &\geq 0, & \forall i, i_2 \in I, & k_1, k_2 \in K.
\end{align*}
\]

Thus, the linearized L1M becomes

**Linearized Level-1 Model (L1M):**

Maximize \( \sum_{i \in I} \sum_{k_1, k_2 \in K} E_{i,k_1} P_{i,k_1} w_{i,k_1,k_2} \)
subject to \( (9), (10), (11), (12), (16), (17), (18), (19) \), and for completeness, we present the linearized versions of the L2M and IM as follows.

**Linearized Level-2 Model (L2M):**

Maximize \( \sum_{i \in I} \sum_{k_1, k_2 \in K} E_{i,k_1} P_{i,k_1} y_{i,k_1,k_2} \)
subject to \( (13), (14), (15), (16), (17), (18), (19), (20), (21), (22), (23) \).

**Linearized Integrated Model (L1M):**

Maximize \( \sum_{i \in I} \sum_{k_1, k_2 \in K} E_{i,k_1} P_{i,k_1} y_{i,k_1,k_2} \)
subject to \( (2), (3), (4), (5), (6), (7), (8), (20), (21), (22), (23) \).

6. Computational results

In this section, we report our experience solving a real instance of a store layout with maximum exposure problem for a major grocery chain in Europe. Although, in practice, we only need to solve a Level-2 problem in most situations as category locations were fixed, we report our experience solving each of IM, L1M and L2M for completeness.

We conduct the computational experiments based on real data for a typical medium-sized store of the grocery chain mentioned in Section 1. As shown in Figures 1 and 2, the problem involves 20 locations/categories (22 nodes including the store entrance and the store exit) at Level-1 and 48 sublocations/subcategories (50 nodes including the store entrance and store exit) at Level-2, respectively. We employ Gurobi (version 7.0.2), an industry-leading mathematical programming solver package to solve our store layout problems directly using its MIQP or MILP solvers. We report results for solving models specified in IM, L1M and L2M, as well as their linearized counterparts L1M, LL1M and LL2M. The
We observe from Table 3 that Gurobi is able to solve the L1M and L2M problems optimally using the hierarchical sequential approach. It takes 35.5 min to solve the L1M and 127.6 min to solve the subsequent L2M. The OFV of the final optimal solution of L2M is 1,492,115.5, which indicates an 11.6% improvement to the feasible solution (with a USOFV of 1,337,105.5) obtained by Gurobi from the IM. Table 3 also clearly shows the computational time improvement from the linearized models LL1M and LL2M. Gurobi takes just 0.1 min to solve LL1M (by obtaining the same OFV of L1M) and 2.8 min to solve LL2M (by obtaining a slightly better OFV of L2M), respectively. This indicates that the linearized versions of the layout decision models, regardless of tactical or strategic levels, can dramatically reduce the computation time by orders of magnitude. This makes the approach of using LL1M and LL2M attractive as a vital tool in making store layout decisions in practice.

Moreover, we observe that the final OFVs from L2M and LL2M are different: 1,492,115.5 versus 1,496,600.0, with the same optimal OFV at 1,497,932.0 from the associated Level-1 models L1M and LL1M. However, their optimality gaps from L2M and LL2M indicate that both solutions are optimal or very close to the optimal solution. This is because in this particular case, L1M and LL1M yield two alternative optimal solutions with the same OFV. Sequentially, L2M and LL2M built upon two different models based on these two Level-1 solutions, and yield two ultimate different optimal solutions at the tactical level. This confirms that our hierarchical decomposition method that constructs the Level-2 model based on the optimal Level-1 solution may lead to sub-optimality for the original problem, and is a heuristic by nature.

Although we emphasize that L2M or LL2M are the models required to make short-term layout decisions needed in most practical applications, there are a few circumstances where combined strategic and tactical layout decisions need to be made simultaneously (e.g., in a store remodelling). Since it is impractical to employ IM or L1M, we recommend the adoption of an approximate method like the two level hierarchical sequential approach described above. As discussed earlier, such an approach of decomposing the original problem into two subproblems then solving them sequentially may lead to a sub-optimal final solution for the original problem. To alleviate this shortcoming, we devise a naive heuristic that simply allows Gurobi to generate a pool of optimal or near optimal solutions at Level-1 model, with the pool size V. Then, for each of the V candidate L1M/LL1M solution in the solution pool, we construct the corresponding Level-2 models and solve them to optimality. Finally, we select the best solution from these V Level-2 solutions as the final best solution. We implement such a mechanism in our two-level hierarchical framework by allowing 10 optimal/near optimal solutions in the solution pool for L1M/LL1M. Note that the solutions in the pool

| Result | Model | IM | LIM |
|--------|-------|----|-----|
| CPU Time (minutes) | 0.1 | 0.1 |
| Optimal Gap (%) | 0.0095 | 0.0094 |
| Solution OFV | 1,492,115.5 | 1,496,600.0 |
| CPU Time (minutes) | 35.5 | 127.6 |
| Optimal Gap (%) | 0.0000 | 0.0000 |
| Solution OFV | 1,497,932.0 | 1,492,115.5 |

**Table 3.** Computational results on the two level models.

| Result | Model | L1M | L2M | LL1M | LL2M |
|--------|-------|-----|-----|------|------|
| CPU Time (minutes) | 35.5 | 127.6 | 0.1 | 2.8 |
| Optimal Gap (%) | 0.0095 | 0.0094 | 0.0000 | 0.0000 |
| Solution OFV | 1,497,932.0 | 1,492,115.5 | 1,497,932.0 | 1,496,600.0 |

**Table 2.** Computational results on the integrated models IM and LIM.
are selected so that the OFVs are all within 0.1% of the best solution. For illustrative purpose, we list OFVs of 10 pooled solutions in LL1M and LL2M in Table 4.

We further summarize the results of our pooled heuristic in Table 5.

Table 4. OFVs for 10 pooled solutions.

| Solution # | L1M      | L2M      | LL1M     | LL2M     |
|------------|----------|----------|----------|----------|
| 1          | 1,147,932.0 | 1,492,115.5 | 1,147,932.0 | 1,496,600.0 |
| 2          | 1,147,932.0 | 1,489,908.5  | 1,147,932.0 | 1,492,115.5 |
| 3          | 1,147,932.0 | 1,498,810.0  | 1,147,932.0 | 1,498,810.0 |
| 4          | 1,147,932.0 | 1,496,600.0  | 1,147,932.0 | 1,491,062.0 |
| 5          | 1,147,932.0 | 1,495,566.5  | 1,147,932.0 | 1,489,908.5 |
| 6          | 1,147,932.0 | 1,491,062.0  | 1,147,932.0 | 1,495,566.5 |
| 7          | 1,147,932.0 | 1,497,756.5  | 1,147,932.0 | 1,488,875.0 |
| 8          | 1,147,932.0 | 1,488,875.0  | 1,147,932.0 | 1,497,756.5 |
| 9          | 1,147,905.0 | 1,486,158.0  | 1,147,881.0 | 1,496,600.0 |
| 10         | 1,147,905.0 | 1,481,837.5  | 1,147,881.0 | 1,498,810.0 |

Table 5. Computational results for pooled heuristic using pool size 10.

| Results/Model | L1M —L2M | LL1M —LL2M |
|---------------|-----------|------------|
| Level-1 Worst/Best OFV | 1,147,905.0/1,147,932.0 | 1,147,881.0/1,147,932.0 |
| Level-2 Worst/Best OFV | 1,481,837.5/1,498,810.0 | 1,488,875.0/1,498,810.0 |
| Final Best OFV | 1,498,810.0 | 1,498,810.0 |
| CPU Time (minutes) | 1428.8 | 40.4 |

The exposure is lower as subcategories are only located on one side. On the other hand, in the optimal layout in Figure 3(b), traffic density is much more distributed as compared to the current layout. Some eminent observations may provide explanations on how total exposure is improved. First, the top selling category, “Breads” (13–16), is moved from right-hand side of the store to the back of the store in the Level-1 optimization. As a result, it leads to more traffic and exposure deeper into the store as customers have to pass by the grid aisles to pick bread products. Second observation comes from “Fresh Dairy Drink” (21), “Leafy Green Vegetables” (6), “Cakes & Cookies” (46) and “Signature Breads” (16), four of the eight top selling subcategories. Subcategories in the first pair, “Fresh Dairy Drink” (21) and “Leafy Green Vegetables” (6), and the second pair “Cakes & Cookies” (46) and “Signature Breads” (16) co-exist in more than 20% of the total baskets. In the current layout, these pairs are located facing each other, therefore, a customer picks subcategories in pairs without getting exposed to other subcategories. In contrast, the optimization exploits these associations between the subcategories and disperses them in the store layout to realize more exposure. In the Level-1 optimization, first, the “Produce” (5–8) category moves to peripheral shelves at the right-hand side of the store away from “Fresh Dairy Drink” (21). Then, in the Level-2 optimization, “Leafy Green Vegetables” (6) moves to the far corner of the store which in return further increases the exposure. Similarly, “Cakes & Cookies” (46) and “Signature Breads” (16), which are located close to the store exit in the current layout, are reasonably spaced out in the optimal layout where the former is moved to the centre of the leftmost grid aisle and the latter is moved to the centre-back of the store. Finally, five out of eight subcategories (28, 31, 34, 40, 46), which are located in the section of grid aisles that is nearer the front of the store in the current layout, are moved to the centre or end of grid aisles in Level-2 optimization. Consequently, traffic concentration at the front of the store in the current layout is dispersed to the store centre along with more exposure in the optimal layout.

We presented a real-life example where customers pick the items in their basket according to their popularity. However, in reality, customers may pick items in a random order, especially, if they are not familiar.
with the item locations. For instance, customers of the stores around tourism hot-spots and business hubs, which attract a large number of visitors, may randomly browse the store and pick the items they need. On the other hand, stores in the residential areas have more loyal customers who are purposefully treading aisles to pick the items in their list. Therefore, it is important to evaluate the sensitivity of the solutions to the format of order picking. We solve the optimization problem for different compositions of order picking formats where percentage of “random” order picks range from 0% to 50% in total orders. For each percentage level, we conduct 20 experiments where we sample 5,000 transactions from total of 20,842 transactions for each experiment. Furthermore, we randomly determine the “random” order picks in each experiment and take the remaining as “popularity” order picks.

Figure 4 shows the boxplots for exposure improvement and total travelled distance with respect to “random” order pick percentages. Experiments suggest that both expected improvement and total travelled distance are moderately sensitive to order picking format, they both increase with the increasing percentage of “random” order picks. This result is expected as customers travel longer and hence are exposed to more subcategories if they pick randomly. Moreover, experiments enable us to derive data-driven insights which can benefit retailers in layout decisions. First, regardless of the “random” order pick percentage, we observe that the top selling category, “Breads” (13–16), is always located at the back of the store in the optimal layout of all the experiments. Secondly, we see that the high selling categories are evenly filling the floor space in all of the optimal layouts. Therefore, retailers could realize gains if high selling categories are dispersed in
the store layout so that customers’ shopping paths cover more of the store and more products are exposed. Finally, associated subcategory pairs tend to be located farther from each other in the optimal layout as “popularity” order pick percentage increases. Traditionally, retailers prefer to locate associated pairs close to each other to preserve shopping convenience and habits, however, they can create significant exposure opportunity if they stretch this business rule and space out associated pairs as much as possible. Thus, the trade-off between the customer convenience and the product exposure should become an important factor to be explored in future store layout decisions.

In summary, our findings suggest that the most popular category should be placed at the back of the store and popular categories should evenly cover the floor space regardless of the store’s customer profile. On the other hand, if the store has a large loyal customer base, retailers can benefit from exploiting shopping habits and increase the exposure by
spacing out associated subcategory pairs in the layout with minimum disturbance to shopper convenience.

7. Conclusion

In this research, we propose a data-driven optimization methodology for retail store layout. This article contributes to the literature with a novel optimization framework which can support layout planning operations over a wide time span along with maximization of total product visibility on the layout through well-depicted customer mobility using shopper paths. Specifically, this new approach optimizes the store layout while conforming to retailer’s layout operation constraints and decisions over entire strategic planning horizon and dynamically accounting for tactical changes in product exposures due to resulting changes in store traffic and customer purchase patterns. Hierarchical optimization models, Level-1 at the strategic level and Level-2 at the tactical level, integrate into decision processes for both long and short terms. Optimization models are not only in-line with decision processes but also satisfy additional practical business constraints such as suitability of fixtures. In both models, category and subcategory assignments to optimal locations maximize overall product exposure by inducing customer shopping paths that maximize traversal of the store layout. Optimization over customer shopping paths provides an edge to the approach as the store automatically adapts its layout to the changes in the customer habits and industry trends.

We formulate the integrated store layout problem at both strategic and tactical levels as a restricted case of the hard-to-solve quadratic assignment problem. Furthermore, we decompose the integrated problem into the two sequential sub-problems at the strategic and tactical levels, respectively. The Adams-Johnson linearization technique is used to expedite the solution time by orders of magnitude. We show that the real world problems at the strategic level as well as the tactical level can be readily solved using a standard MILP solver package like Gurobi. In addition, the pooled heuristic coupled with the linearization technique is proven effective for solving the integrated model when needed. Results using the data set from a real grocery store clearly show the benefit in using the approach as total exposure is significantly increased while satisfying operational requirements.

The research described in this article reflects our initial effort to address the retail store layout decisions based on data-driven analytics and optimization methods. The problem set up in this article provides a baseline for retail store layout optimization on which several extensions of the problem can consider. First, we assume categories and subcategories can be interchanged because they have equal size. As an extension, problem can be solved by addressing complications of size and interchangeability constraints and different store layouts. Secondly, the problem can be transformed into an impulse revenue maximization problem by converting product exposures to revenues using item level impulse purchase rates or expert opinion as discussed in the literature review. Another natural extension would be to combine the problem with the shelf/floor space allocation problem. We believe the latter two extensions would be more meaningful with rigorous estimation of impulse revenues.

As a final note, the methodology proposed can be flexibly configured to serve different business objectives. Retailers desire similar layouts for stores in a region to retain familiarity and easy navigation as customers may visit different stores. In-line with this requirement, first, the strategic Level-1 problem can be solved for a group of stores with similar layouts for pooled shopper paths. The resulting optimal layout would both provide the uniformity and maximize the total exposure for all stores. Next, each store’s layout can be customarily optimized considering their own shopper paths which further boost product exposure without harming the store identity.

Acknowledgments

We sincerely thank the helpful feedback from the editor, associate editor, and two anonymous referees, that helped significantly improve the clarity, presentation and quality of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Adams, W. P., & Johnson, T. A. (1994). Improved linear programming-based lower bounds for the quadratic assignment problem. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 16, 43–77.

Ainsworth, J., & Foster, J. (2017). Comfort in brick and mortar shopping experiences: Examining antecedents and consequences of comfortable retail experiences. *Journal of Retailing and Consumer Services*, 35, 27–35.

Baker, J., Levy, M., & Grewal, D. (1992). An experimental approach to making retail store environmental decisions. *Journal of Retailing*, 68(4), 445.

Ballester, N., Guthrie, B., Martens, S., Mowrey, C., Parikh, P. J., Zhang, X. (2014). Effect of retail layout on traffic density and travel distance. In *IIE Annual Conference Proceedings* (p. 798).
Bellenger, D. H., Robertson, D. H., & Hirschman, E. C. (1978). Impulse buying varies by product. *Journal of Advertising Research*, 6(18), 15–18.

Boros, P., Feher, O., Lakner, Z., Niroomand, S., & Vizvari, B. (2016). Modeling supermarket re-layout from the owner’s perspective. *Annals of Operations Research*, 238(1–2), 27–40.

Botsali, A. R. (2007). *Retail facility layout design* [Unpublished doctoral dissertation]. Texas A&M University.

Botsali, A. R., Peters, B. A. (2005). A network based layout design model for retail stores. In *IIE Annual Conference. Proceedings* (p. 1).

Brixius, N. W. (2000). *Solving large-scale quadratic assignment problems* [Unpublished doctoral dissertation]. The University of Iowa.

Cil, I. (2012). Consumption universes based supermarket layout through association rule mining and multidimensional scaling. *Expert Systems with Applications*, 39(10), 8611–8625.

Drez, X., Hoch, S. J., & Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing*, 70(4), 301–326.

Flamand, T., Ghoniem, A., & Maddah, B. (2016). Promoting impulse buying by allocating retail shelf space to grouped product categories. *Journal of the Operational Research Society*, 67(7), 953–969.

Flamand, T., Ghoniem, A., Haouari, M., & Maddah, B. (2018). Integrated assortment planning and store-wide shelf space allocation: An optimization-based approach. *Omega*, 81, 134–149.

Frieze, A. M., & Yadegar, J. (1983). On the quadratic assignment problem. *Discrete Applied Mathematics*, 5(1), 89–98.

Glover, F., & Woolsey, E. (1974). Converting the 0–1 polynomial programming problem to a 0–1 linear program. *Operations Research*, 22(1), 180–182.

Guthrie, B., & Parikh, P. J. (2020). The rack orientation and curvature problem for retailers. *IIE Transactions*, 52(10), 1081–1097.

Guthrie, B., & Parikh, P. J. (2021). Evaluating exposure of a retail rack layout in 3D. *Flexible Services and Manufacturing Journal*, 33(1), 107–135.

Hirpara, S., & Parikh, P. J. (2021). Retail facility layout considering shopper path. *Computers & Industrial Engineering*, 154, 106919.

Hui, S. K., Inman, J. J., Huang, Y., & Suher, J. (2013). The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing*, 77(2), 1–16.

Inman, J. J., Winer, R. S., & Ferraro, R. (2009). The interplay among category characteristics, customer characteristics, and customer activities on in-store decision making. *Journal of Marketing*, 73(5), 19–29.

Karki, U., Guthrie, B., & Parikh, P. J. (2021). Joint determination of rack configuration and shelf space allocation for a retailer. *International Journal of Production Economics*, 234, 107943. https://doi.org/10.1016/j.ijpe.2020.107943.

Kaufman, L., & Broeckx, F. (1978). An algorithm for the quadratic assignment problem using bender’s decomposition. *European Journal of Operational Research*, 2(3), 207–211.

Kholod, M., Nakahara, T., Azuma, H., Yada, K. (2010). The influence of shopping path length on purchase behaviour in grocery store. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 273–280).

Kollat, D. T., & Willett, R. P. (1967). Customer impulse purchasing behaviour. *Journal of Marketing Research*, 4(1), 21–31.

Koopmans, T. C., & Beckmann, M. (1957). Assignment problems and the location of economic activities. *Econometrica: Journal of the Econometric Society*, 25(1), 53–76.

Larson, J. S., Bradlow, E. T., & Fader, P. S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing*, 22(4), 395–414.

Lawler, E. L. (1963). The quadratic assignment problem. *Management Science*, 9(4), 586–599.

Levy, M., Weitz, B. A., Grewal, D., & Madore, M. (2004). *Retail management*. McGraw-Hill Irwin New York.

Lewison, D. M. (1994). *Retailing*. Macmillan College Publishing Company.

Merrilees, B., & Miller, D. (2001). Superstore interactivity: A new self-service paradigm of retail service? *International Journal of Retail & Distribution Management*, 29(8), 379–389.

Mishra, A., & Mishra, H. (2010). We are what we consume: The influence of food consumption on impulsive choice. *Journal of Marketing Research*, 47(6), 1129–1137.

Mowrey, C. H., Parikh, P. J., & Gue, K. R. (2018). A model to optimize rack layout in a retail store. *European Journal of Operational Research*, 271 1100–1112.

Mowrey, C. H., Parikh, P. J., & Gue, K. R. (2019). The impact of rack layout on visual experience in a retail store. *INFOR: Information Systems and Operational Research*, 57(1), 75–98.

Ozgormus, E., & Smith, A. E. (2020). A data-driven approach to grocery store block layout. *Computers & Industrial Engineering*, 139, 105562.

Peng, C.-C. (2011). Optimizing the layout of a grocery/convenience store to maximize revenue from impulse items [Unpublished doctoral dissertation]. State University of New York at Buffalo.

Rook, D. W. (1987). The buying impulse. *Journal of Consumer Research*, 14(2), 189–199.

Sahni, S., & Gonzalez, T. (1976). P-complete approximation problems. *Journal of the ACM*, 23(3), 555–565.

Vrechopoulos, A. P., O’keefe, R. M., Doukidis, G. I., & Siomkos, G. J. (2004). Virtual store layout: An experimental comparison in the context of grocery retail. *Journal of Retailing*, 80(1), 13–22.

Watters, L. J. (1967). Reduction of integer polynomial programming problems to zero-one linear programming problems. *Operations Research*, 15(6), 1171–1174.

West, J. C. (1951). Results of two years of study into impulse buying. *Journal of Marketing*, 3(15), 362–363.

Xu, J., Gul, E., & Lim, A. (2021). Maximizing store revenues using tabu search for floor space optimization. *International Journal of Revenue Management*, 12(1–2), 56–82 https://doi.org/10.1504/IJRM.2021.114969.

Yapiciglu, H., & Smith, A. E. (2012a). A bi-objective model for the retail spatial design problem. *Engineering Optimization*, 44(3), 243–266.

Yapiciglu, H., & Smith, A. E. (2012b). Retail space design considering revenue and adjacencies using a racetrack aisle network. *IIE Transactions*, 44(6), 446–458.