

CAPTURING DEMAND TRANSFERENCE IN RETAIL - A STATISTICAL APPROACH

Omker Mahalanobish

Statistical Analyst, Walmart Labs, Bengaluru, India
omker.mahalanobish@walmart.com

Souraj Mishra

Statistical Analyst, Walmart Labs, Bengaluru, India
souraj.mishra@walmart.com

Amlan Das

Statistical Analyst, Walmart Labs, Bengaluru, India
amlan.das@walmart.com

Subhasish Misra*

Associate Data Scientist, Walmart Labs, Bengaluru, India
subhasish.misra@walmart.com

Background:

While an item substitution measure provides for the direction, **demand transference quantifies the magnitude of demand** that may get transferred to an item a) When its substitute is deleted b) When it is introduced in a store and cannibalizes on similar items.

This, hence, is an important input into assortment optimization. If an item is predicted to exhibit a good extent of transference **then we may be more certain of deleting it** (provided, it is less than an average performer in terms of sales). Conversely, we should be careful of deleting a very incremental item (with low demand transference) – since we'll be losing on a bulk of its demand.

Note that transference is not explicitly observed, it's latent. Our methodology explains how we capture it.

Method:

Data: POS, promotions & item attribute data is harnessed for this process.

Modeling:

- Regression models (in a longitudinal setup) are used to estimate demand for an item – among other explanatory variables we have one that accounts for cannibalization effect of similar items.
- The cannibalization term uses the attribute data to calculate item similarity. Its value changes depending on presence/absence of similar items and is the instrument through which demand transference seeps into this model.

- The modeling process is designed to automatically take care of complications such as multicollinearity and sundry regression violations.
- Since each store is unique in terms of the consumer demand pattern these models have been estimated at a store x substitutable community level.
- This means that for a category with 10 + substitutable community, we are estimating $10 * 4000 + = 40000 +$ models using parallelization techniques in Hadoop.

In conclusion, these models predict the extent of transference (i.e. if an item “i₁” in the pre-delete scenario was selling 100 units, then what amount of its demand would get transferred to its substitutes, say, “i₂”, “i₃”, “i₄”). All this, at an individual store level as well as the overall US.

Expected outcome:

The methodology has been successfully tested for multiple foods and consumable categories, as well as general merchandising categories in the US – efforts are on towards making this one of the processes of estimating demand transference. The entire process, despite involving sophisticated modeling has been **scaled (across all stores), automated and productized** as an easy to use manner for the business user.

Keywords: *Regression, Cannibalization, Retail, Parallelization, Forecasting*

1. INTRODUCTION

Assortment is a key element of a retailer’s marketing mix. It differentiates a retailer from its competitors and has a very strong influence on retail sales. Retailers face the problem of selecting the assortment that maximizes category profitability, without sacrificing customer satisfaction.

Although some headway has been made in the context of assortment optimization, practitioners and academics agree that more research is needed to provide feasible solutions to realistic assortment problems. Specifically, the challenge of assortment optimization is compounded by the fact that the demand for an item cannot be assumed to be fixed; it is instead affected by the presence of other items as a result of product substitution.

One of the important challenge is to account for similarity effects: an item is a stronger substitute for similar items than it is for dissimilar items. Demand is also driven by own- and cross-marketing mix instruments such as price, promotion and by heterogeneous preference across stores. Capturing these aspects in a response model is further complicated by the fact that assortments and prices observed in empirical data are unlikely to be exogenous. Finally, retailers have to decide about not only the assortment, but also about the pricing, and these decisions need to be customized at a store level.

In the process of optimizing the store assortment, it is important to understand the process of demand transference. Demand transference is defined as the process of transfer of demand among the items in a store, once a change in assortment is realized.

In a store, for a given category, there may be two realizations of an assortment change :

1. When one or more items are dropped from the assortment, customers who intended to buy any of the dropped items, might either choose to opt for another ‘substitutable’ item or walk away from the store, without a purchase.

2. When one or more items are introduced into the assortment, customers who purchased any of the new items, might either buy the new item out of impulse or replace purchasing an existing item with the new item.

A better understanding of this underlying process would help in identifying the optimum assortment for the particular category in the particular store. In this paper, we aim to model the mechanism of demand transference so as to optimize the store assortment, for the category.

2. LITERATURE REVIEW

This section briefly discusses about the studies in place related to assortment selection / optimization.

The common points among the available literatures is that they all look to optimize the assortment, based on maximizing cost function (usually sales or profit). We here would only restrict ourselves towards those studies which deal with the item attributes along with the scanner data.

Among the available articles, Fisher and Vaidyanathan (Fisher et al., 2009) look into selecting the optimum number of items from the available lot, to maximize sales. They have defined an approach, in which they view a item as a set of attribute values, use sales history of the items currently carried by the retailer to estimate the demand for each of the attribute values and from this, the demand for any potential item currently not carried by the retailer. They also introduce a model of substitution behavior, estimate the parameters and consider the impact of substitution in choosing assortment.

Kök and Fisher (2007), also tread similar lines wherein they study an assortment planning model in which consumers might accept substitutes when their favorite product is unavailable. They develop an algorithmic process to help retailers compute the best assortment for each store, by estimating the parameters of substitution behavior and demand for products at a store, including products that have not been carried previously in that store. Finally, they propose an iterative optimization heuristic to solve the assortment planning problem.

Other articles like Roederkerk et al., (2013) look into price optimization along with promotion and shelf space optimization. Herein, they adopt a scalable assortment optimization method that allow for theory based substitution patterns and cross marketing mix effects. For the optimization part, they propose a large neighborhood search heuristic methodology.

Our study though on similar lines, addresses an entirely different aspect of assortment study. This is more of scenario based, to understand how the store assortment performs when a change in the store assortment is realized. Basically, we help the retailer to decide which items to drop from the assortment, by helping him understand where the demand of the deleted item would flow to and in what magnitude. The retailer also gets a glimpse of additional incremental demand as well as the magnitude and direction of cannibalization that might be realized when additional items are added to the assortment. In the process of obtaining these insights, he also gets an understanding as to how the items in the new assortment will perform in the future.

3. METHODOLOGY

Demand is a latent feature, which can be experienced but not explicitly observed. Thus, modeling of demand and validation of the same becomes difficult. The nearest proxy to demand is sales. So, here we try to model the sales of each of the products offered in an assortment.

Our methodology consists of a sales model, described in section 3.1 and a predictive algorithm based on the sales model, as described in section 3.2.

3.1. Sales model

Before formulating the model, we look at a significant challenge encountered while developing the sales model: modeling store level UPC sales on attributes. UPC (Universal Product Code) is used to identify trade items in stores across different retailers and markets. UPC is aggregation of items which can vary across different regions.

3.1.1. Modeling framework

To model UPC sales at a store level, we chose store-level scanner data, as it provides a holistic view of the available assortment in the store. An example of the input data is provided in Table 1 under section 4.1. Our approach of modeling UPC sales on UPC attributes is motivated by the assertion that customers do not form preference of each individual UPC in a product category but that these preferences are derived from the preferences for the underlying attributes (e.g., size, brand, flavor, etc.). Theoretical justification of the same is available in economics (Lancaster, 1971) and psychology (Fishbein, 1967).

Our model thus takes into account the UPC attributes, in order to model UPC sales. Apart from the UPCs own attributes, attributes of other available UPCs would also affect the sales. We thus incorporate variables to account for a UPCs attribute similarity as well as cross attribute similarity with the other UPCs in the assortment.

3.1.2. Modeling formulation

We would now develop the attribute based model and highlight the role similarity variables play. While modeling UPC sales at a store level, we allow for flexible substitution patterns, and non-linear effects by starting with a log-log model (Roederkerk et al., 2013), similar to the SCAN*PRO model (Wittink et al., 1988):

$$\log(S_{kti}) = \underbrace{\alpha_{ki}}_A + \underbrace{\beta \cdot \log(P_{kti})}_B + \underbrace{\sum_{m \in \mathcal{A}} \gamma_{kmti}}_C \quad (1)$$

where,

S_{kti} = unit sales of UPC $k \in \{1, 2, \dots, K\}$ in week $t \in \{1, 2, \dots, T\}$ in store $i \in \{1, 2, \dots, n\}$;

α_{ki} = UPC-store intercept for UPC $k \in \{1, 2, \dots, K\}$ in store $i \in \{1, 2, \dots, n\}$;

P_{kti} = price of UPC $k \in \{1, 2, \dots, K\}$ in week $t \in \{1, 2, \dots, T\}$ in store $i \in \{1, 2, \dots, n\}$;

γ_{kmti} = similarity score of UPC $k \in \{1, 2, \dots, K\}$, for attribute $m \in \mathcal{A}$, in week $t \in \{1, 2, \dots, T\}$ in store $i \in \{1, 2, \dots, n\}$;

\mathcal{A} = set of all attributes, evaluated for all UPCs in a product category;

Further, α_{ki} may be replaced by strictly store level intercepts along with attribute dummies such that

$$\alpha_{ki} = \underbrace{\alpha_i}_A + \underbrace{\sum_{m \in \mathcal{A}} \sum_{l=1}^{m_l} A_{kml}}_B \quad (2)$$

where,

A_{kml} = 1 if UPC k possesses level l of attribute $m \in \mathcal{A}$, and 0 otherwise, if m is nominal

A_{kml} = the realization of attribute $m \in \mathcal{A}$, if m is metric

3.1.3. Attribute similarity score

The similarity score of UPC k , for a nominal or metric attribute m , in week t , in store i is defined such that it varies between 0 (minimum similarity) and 1 (maximum similarity), and also reflects the similarity of UPC k relative to the distribution of attribute m in the entire available assortment.

Let $SIM_{kk'mti}$ denote the magnitude of similarity between UPC k and UPC k' with respect to attribute m in store i in week t .

Further to the above discussed features of similarity, if UPC k and UPC k' share the same level of nominal attribute m , then the perceived similarity of UPC k and UPC k' should be stronger when their shared attribute level occurs less frequently (Goodall, 1966). We obtain all the above, by defining :

$$SIM_{kk'mti} = I\{A_{k'm} = A_{km}\} \cdot \left(1 - \frac{1}{N_{ti}} \cdot \sum_{\substack{k''=1 \\ x_{k''ti}=1}}^K I\{A_{k''m} = A_{km}\}\right) \quad (3a)$$

if attribute m is nominal, where,

$I\{.\}$ = an indicator function which takes the value 1 if its argument holds, 0 otherwise;

A_{km} = the level attained by UPC k on attribute m such that $A_{km} = l \Leftrightarrow A_{kml} = 1$;

N_{ti} = the number of UPCs present in week t in store i ;

$x_{kti} = 1$, if UPC k was available in store i , for at least 1 day in week t ; else 0.

Table 3 in section 4.2 illustrates how this would work for a Brand attribute.

On the other hand, the similarity of UPC k and UPC k' , with respect to a metric attribute m , is perceived to be more if there exists fewer UPCs with attribute values between the attribute values of UPC k and UPC k' . This is obtained by defining:

$$SIM_{kk'mti} = 1 - \frac{1}{N_{ti}} \cdot \sum_{\substack{k''=1 \\ x_{k''ti}=1}}^K I\{\min(A_{km}, A_{k'm}) \leq A_{k''m} \leq \max(A_{km}, A_{k'm})\} \quad (3b)$$

if attribute m is metric.

This definition is numerically illustrated for Weight attribute in Table 5 in section 4.2.

Once we have described the measure of similarity for UPC k and UPC k' , we may now formulate the similarity score of UPC k for attribute m in week t in store i as :

$$\gamma_{kmti} = \text{mean}^*_{k' \neq k}(SIM_{kk'mti}) \quad (4)$$

where,

$\text{mean}^*(.)$ = Arithmetic Mean of the non-zero elements of the argument, if attribute m is nominal, usual Arithmetic Mean otherwise.

3.1.4. Model implementation

The model described in this paper, is best implemented when modeled category wise. Now, each category has properties of its own and consists of widely different varieties of UPCs. The two major category properties that is observed is as follows :

1. Demand might get transferred to any and every UPC of the category and

2. Transfer of demand is restricted only within mutually exclusive and exhaustive set of substitutable groups, which are very different from each other (further discussed in section 4.3).

For case 2, one may carry on the same analysis over each substitutable group, as if assuming it to be a small sized category of sorts.

Since we have formulated a linear regression as mentioned in (1), all regression sanity checks have been taken care of and the final model thus only consists of the uncorrelated and significant regressors among the ones mentioned in (1).

3.2. Predictive algorithm

We would now look into how to predict the magnitude of demand transference and the walkoff rate.

Define

\mathbb{A}_i : the training assortment of store i ;

\mathbb{A}'_i : the assortment of store i after the assortment change;

Now, for every UPC in \mathbb{A}_i , we can easily obtain the predicted weekly unit sales from the model as explained in (1).

Also, the values of parts A and B in (1) are independent of the store assortment (assuming there is no change in price in any of the items in \mathbb{A}_i) and thus doesn't change. It suffices to compute these values only for those UPCs that have been introduced in \mathbb{A}'_i but were not a part of \mathbb{A}_i . Part C in (1) directly depends upon the current assortment in store and hence the similarity score is recalculated for each UPC in the new assortment. Once we have all the required information, the predicted weekly unit sales of every UPC in \mathbb{A}_i can be easily obtained.

Define,

\hat{S}_{ki} = predicted weekly unit sales of UPC $k \in \mathbb{A}_i$;

\hat{S}'_{ki} = predicted weekly unit sales of UPC $k \in \mathbb{A}'_i$;

Therefore,

$\Delta S_{ki} = \hat{S}'_{ki} - \hat{S}_{ki}$, is the change in the weekly unit sales of UPC $k \in \mathbb{A}_i \cap \mathbb{A}'_i$.

But, $\Delta S_{ki} = \hat{S}'_{ki}$, if UPC $k \in \mathbb{A}'_i \setminus \mathbb{A}_i$

3.2.1. Case of item deletion

Define \mathcal{U}_{del} = set of UPCs that have been deleted from \mathbb{A}_i , and are not present in \mathbb{A}'_i . Then,

$$\Delta_{k\mathbb{A}'_i}^{del} = \frac{\Delta S_{ki}}{\sum_{t \in \mathcal{U}_{del}} \hat{S}_{ti}} \cdot 100 \% \quad (5a)$$

where,

$\Delta_{k\mathbb{A}'_i}^{del}$ = demand of UPCs in \mathcal{U}_{del} transferred to UPC k , $\forall k \in \mathbb{A}'_i$.

Herein, the walk-off rate is calculated as:

$$\omega_{\mathbb{A}'_i}^{del} = 100 - \sum_{k \in \mathbb{A}'_i} \Delta_{k\mathbb{A}'_i}^{del} \quad (5b)$$

3.2.2. Case of item addition

Define, \mathcal{U}_{add} = set of UPCs that have been added to \mathbb{A}'_i , but were not a part of \mathbb{A}_i . Then,

$$\Delta_{k\mathbb{A}'_i}^{add} = \frac{|\Delta S_{ki}|}{\sum_{t \in \mathcal{U}_{add}} \hat{S}'_{ti}} \cdot 100 \% \quad (6a)$$

where,

$\Delta_{k\mathbb{A}'_i}^{add}$ = demand of UPCs in \mathcal{U}_{add} cannibalized from UPC k , $\forall k \in \mathbb{A}'_i$.

Herein, the incrementality is calculated as:

$$\omega_{\mathbb{A}'_i}^{add} = 100 - \sum_{k \in \mathbb{A}'_i} \Delta_{k\mathbb{A}'_i}^{add} \quad (6b)$$

3.2.3. Case of both item deletion and item addition

In this case, it becomes difficult to identify separately, what amount of the change in the demand for UPC k , is due to the transfer of demand from the deleted UPCs and how much amount is due to cannibalization of the added UPCs. Further, there could even be some amount of demand transference towards the newly added UPCs as well.

Hence, one may separately consider the deletions and additions to obtain the demand transference measures.

Therefore, we have

$$\Delta_{k\mathbb{A}'_i}^{add} = \frac{|\Delta S_{ki}| \cdot \frac{|\mathcal{U}_{add}|}{|\mathcal{U}_{del}|}}{\sum_{t \in \mathcal{U}_{add}} \hat{S}'_{ti}} \cdot 100 \%, \quad \forall k \in \mathbb{A}'_i \quad (7a)$$

$$\Delta_{k\mathbb{A}'_i}^{del} = \frac{\Delta S_{ki} \cdot \frac{|\mathcal{U}_{del}|}{|\mathcal{U}_{add}|}}{\sum_{t \in \mathcal{U}_{del}} \hat{S}_{ti}} \cdot 100 \%, \quad \forall k \in \mathbb{A}'_i \quad (7b)$$

where,

$\Delta_{k\mathbb{A}'_i}^{add}$ = demand of UPCs in \mathcal{U}_{add} cannibalized from UPC k , $\forall k \in \mathbb{A}'_i$.

$\Delta_{k\mathbb{A}'_i}^{del}$ = demand of UPCs in \mathcal{U}_{del} transferred to UPC k , $\forall k \in \mathbb{A}'_i$.

In (7b),

$$\Delta S_{ki} = \Delta S_{ki} \cdot \left(1 - \sum_{t \in \mathcal{U}_{add}} \frac{\Delta_{t\mathbb{A}'_i}^{add}}{100} \cdot I\{k \in \mathcal{U}_{add}\}\right) \quad (7c)$$

Therefore, walkoff rate is:

$$\omega_{\mathbb{A}'_i}^{del} = 100 - \sum_{k \in \mathbb{A}'_i} \Delta_{k\mathbb{A}'_i}^{del} \quad (7d)$$

and incrementality is defined as:

$$\omega_{\mathbb{A}'_i}^{add} = 100 - \sum_{k \in \mathbb{A}'_i} \Delta_{k\mathbb{A}'_i}^{add} - \sum_{k \in \mathcal{U}_{add}} \Delta_{k\mathbb{A}'_i}^{del} \quad (7e)$$

4. DISCUSSION

Here, we will have a brief walkthrough of a sample input data for the algorithm along with the similarity calculation for the same.

4.1. Scanner data and attribute data

Table 1 refers to a snapshot of the scanner data that we require to carry on with the analysis. The snapshot here has been restricted to 4 UPCs, in 3 weeks and 1 store.

Table 1 Snapshot of the scanner data

Store No.	UPC	Week No.	Units sold	Dollar Sales	Price	Days available
1	UPC 1	1	2	2.40	1.20	7
1	UPC 1	2	3	3.55	1.18	7
1	UPC 1	3	2	2.40	1.20	7
1	UPC 2	1	6	4.50	0.75	6
1	UPC 2	2	7	5.25	0.75	7
1	UPC 2	3	2	1.60	0.80	7
1	UPC 3	1	0	0.00	-	0
1	UPC 3	2	3	4.50	1.50	3
1	UPC 3	3	1	1.50	1.50	4
1	UPC 4	1	10	6.00	0.60	7
1	UPC 4	2	8	4.80	0.60	7
1	UPC 4	3	2	1.24	0.62	2

Table 2 Attribute information for UPCs in the snapshot

UPC	Brand	Weight (in gm)
UPC 1	Brand 1	200
UPC 2	Brand 1	180
UPC 3	Brand 1	200
UPC 4	Brand 2	150

4.2. Computing the attribute similarity score

As depicted in Table 2, there are two attributes to take care of viz. Brand (a nominal attribute) and Weight (a metric attribute).

For attribute Brand, Brand 1 is present in 75% of the overall assortment, whereas Brand 2 is present in 25% of the overall assortment.

Hence, according to this example, the similarity scores for UPC 1 with respect to the nominal attribute Brand is demonstrated in Table 3 below:

Table 3 Week wise brand similarity score for UPC 1

Week No.	Brand 1 presence	Brand similarity score (γ_{kmti})
1	66.67 %	0.33
2	75.00%	0.25
3	75.00%	0.25

Similarly, for the metric attribute Weight, similarity score of UPC 1 is seen to be as described in Table 5 below:

Table 4 Weekly weight proximity percent for each UPC

UPC	Week No.	Weight proximity percent	Weight similarity score
UPC 2	1	66.67 %	0.33
UPC 2	2	75.00 %	0.25
UPC 2	3	75.00 %	0.25
UPC 3	1	-	-
UPC 3	2	75.00 %	0.25
UPC 3	3	75.00 %	0.25
UPC 4	1	100.00 %	0.00
UPC 4	2	100.00 %	0.00
UPC 4	3	100.00 %	0.00

Therefore,

Table 5 Weekly weight similarity score for UPC 1

Week No.	Weight Similarity score
1	0.165
2	0.250
3	0.250

4.3. Substitutable groups

A category can be divided into mutually exclusive and exhaustive groups of items, called substitutable groups. A substitutable group consists of items that are more likely to be substitutes of each other, than that of items in the other substitutable groups.

We accomplish this segmentation into substitutable group by using a proprietary graph partition based algorithm.

When, implemented for a substitutable group, the demand transfer algorithm restricts the transfer of demand within the same group; since by definition, there is very less probability of items in other groups to be proper substitutes.

4.4. Parallelization techniques

The entire algorithm was executed in R, for a category with 7 substitutable groups, available in 4500 stores.

While *Hadoop streaming* was used to execute the algorithm over stores; for a store, the *mclapply* function (which uses forking technique) from *parallel* package was used to parallelize over substitutable groups.

For a fixed store, the runtime in R (using forking via *mclapply*) is comparable to the runtime when executed in Python without any scaling up technique.

4.5. Results and success stories

This algorithm has been run for a variety of categories, both General Merchandise and Fast-Moving Consumer Goods (like Yogurt, Light Bulbs, Dish Soap, Utility Pants, Food Storage, etc.) and has been seen to be performing really well.

The Mean Absolute Percentage Error for the model, when validated against observed assortment changes for the aforementioned categories, was almost always in the range of 4% to 13%.

5. CONCLUSION

The problem of demand transference is an important one for any retailer. Obviously, the retailer cannot keep on carrying the same assortment over time. Market trends as well as the item performance, will always compel him to offer his customers the best assortment so as to maximize sales and customer satisfaction. Hence, it is better off to know from beforehand the magnitude of demand transference or cannibalization, that might be experienced with regards to a particular change in his assortment. Having a good understanding of the different scenarios will surely let him plan better than his competitors, and establish his stand in the market.

Wrong choice of item deletion, might have severe repercussions in the form of:

1. churning of customer base which were loyal to the deleted product; or
2. churning of customer base, due to unavailability of proper substitutes of the deleted product in the available assortment.

Similarly, wrong choice of item addition could also be detrimental in the form of the new item not attracting any incremental demand of its own, but is only cannibalizing the demand of the other available items in the assortment, thus not doing any significant good to the retailer.

This study has been aimed to help the retailer address these basic problems of assortment.

6. REFERENCES

- Fishbein M, ed. (1967) *Attitude and Prediction of Behavior* (John Wiley & Sons, New York).
- Fisher ML, Vaidyanathan R (2009) An Algorithm and Demand Estimation Procedure for Retail Assortment Optimization. The Wharton School, Philadelphia, Pennsylvania.
- Goodall DW (1966) A new similarity index based on probability. *Biometrics* 22(4):882–907.

- Kök, G., M. L. Fisher. 2007. Demand Estimation and Assortment Optimization under Substitution: Methodology and Application. *Operations Research* 55(6) 1001–1021.
- Lancaster K (1971) *Consumer Demand: A New Approach* (Columbia University Press, New York).
- Rooderkerk RP, van Heerde HJ, Bijmolt TH (2011) Optimizing Retail Assortments. *Marketing Science* 32(5):699–715.
- Wittink DR, Addona MJ, Hawkes W, Porter JC (1988) SCAN*PRO: The estimation, validation, and use of promotional effects based on scanner data. Working paper, Cornell University, Ithaca, NY.